An Roinn Sláinte
Department of Health

Emergency Medicine Early Warning System (EMEWS)

National Clinical Guideline No. 18

Annex 1: Systematic Review

National Patient Safety Office
Oifig Náisiúnta um Shábháilteacht Othar

NATIONAL CLINICAL EFFECTIVENESS COMMITTEE

NUI Galway
OÉ Gaillimh

Acknowledgments
The research team commissioned by the Department of Health, Clinical Effectiveness Unit, undertook the work described in this report. We thank the Steering Group for this project for their insight and support through the conduct of this work.

We thank Sinead Duane for assistance with initial screening of citations, and Rachel Lee for assisting with inputting citations from additional resources.

This report should be cited as:
Wuytack F, Meskell P, Conway A, McDaid F, Santesso N, Hickey F, Gillespie P, Smith V, Devane D. (2016) Clinical and cost-effectiveness of physiologically based early warning or track and trigger or scoring systems after triage in adult patients presenting to emergency departments: A systematic review. National Clinical Effectiveness Committee, Department of Health: Dublin.

Published by:
The Department of Health
Block 1, Miesian Plaza
50-58 Lower Baggot Street
Dublin 2
D02 XW14
www.health.gov.ie
ISSN 2009-6259

Clinical effectiveness and cost-effectiveness of physiologically based early warning or track and trigger or scoring systems after triage in adult patients presenting to emergency departments: A systematic review

Authors

Francesca Wuytack,[1] Pauline Meskell,[1] Aislinn Conway,[1] Fiona McDaid,[2,3] Nancy Santesso,[4] Fergal Hickey,[5] Paddy Gillespie,[1] Valerie Smith,[1] Declan Devane.[1]

[1] National University of Ireland Galway, Ireland.
[2] Naas General Hospital, Ireland.
[3] National Emergency Medicine Programme, Ireland.
[4] McMaster University, Canada.
[5] Sligo University Hospital, Ireland.

National Clinical Effectiveness Committee (NCEC)

Clinical effectiveness is a key component of patient safety and quality. The integration of best evidence in service provision, through clinical effectiveness processes, promotes healthcare that is up to date, effective and consistent.

The National Clinical Effectiveness Committee (NCEC) is a Ministerial committee established in 2010 as part of the Patient Safety First Initiative. The NCEC is supported by the Clinical Effectiveness Unit (CEU), Department of Health. The NCEC is a partnership between key stakeholders in patient safety and its mission is to provide a framework for national endorsement of clinical guidelines and audit to optimise patient and service user care.

In December 2013, the first National Clinical Guideline (NCG) was published. This was NCEC National Clinical Guideline No. 1 National Early Warning Score (NEWS). It relates to the situation in an acute hospital setting where an adult patient's physiological condition is deteriorating. It was updated in August 2014 to ensure alignment with NCG No. 6 Sepsis Management.

Invitations to tender were issued in July 2015 and a public procurement competition held for the provision of systematic literature reviews and budget impact analysis to support the development of National Clinical Guidelines. Subsequently, a series of reports were commissioned by the CEU/NCEC Department of Health. This report is the first published under this contract. It supports the development of a National Clinical Guideline on Emergency Medicine Early Warning System (EMEWS). A guideline proposal was submitted to the NCEC by the HSE National Clinical Programme for Emergency Medicine and was prioritised for development as a National Clinical Guideline in September 2015.

The Emergency Medicine Early Warning System (EMEWS) is part of a suite of National Clinical Guidelines on Clinical Deterioration. The suite currently consists of:

NCG No	Title	Date
NCG No. 1	National Early Warning Score (NEWS)	February 2013 with clinical update August 2014 Currently being updated.
NCG No. 4	Maternity Early Warning Score (IMEWS)	November 2014
NCG No. 6	Sepsis Management	November 2014 with NICE accreditation Mar 2015
NCG No. 12	Paediatric Early Warning Score system (PEWS)	November 2015
	Emergency Medicine Early Warning System (EMEWS) (Note this was previously known as "Emergency Department Monitoring and Clinical Escalation tool for adults")	Prioritised by the NCEC in September 2015 and development supported by this report.

Further information on the NCEC and National Clinical Guidelines is available at www.health.gov.ie/patient-safety/ncec

Abstract

Background

Changes to physiological parameters precede deterioration of ill patients. Early warning and track and trigger systems (TTS) use routine physiological measurements with pre-specified thresholds to identify deteriorating patients and trigger appropriate and timely escalation of care. Patients presenting to the emergency department (ED) are undiagnosed, undifferentiated and of varying acuity, yet the effectiveness and cost-effectiveness of using early warning systems and TTS in this setting is unclear.

Aim

To provide a rapid systematic review of the evidence of the clinical and cost-effectiveness of physiologically based early warning systems and TTS for the detection of deterioration (post-triage) in adult patients presenting to ED.

Search methods

A comprehensive search of published and unpublished literature, including scientific databases and grey literature resources was carried out. No time filter was used but a filter to include adult patients was applied. No language filter was used, but only information available in English was included.

Selection criteria

Participants were ED adult patients, post-triage. Only early warning systems and TTS that included routine physiological parameters were included. Studies were classified as: (1) Descriptive studies – type and extent of use; (2) Descriptive studies – educational programmes; (3) Guidelines; (4) Effectiveness studies; (5) Development and/or validation studies; and (6) Health economics studies.

Data collection, analysis and quality assessment

Two reviewers independently screened search results by title/abstract and full-text. Data extraction was done by one reviewer with independent verification checks of 50% of records by a second reviewer. Two reviewers conducted quality assessment independently. Data are presented in evidence tables.

Main results

A total of 6397 citations were identified, of which 47 studies, 3 guidelines and 1 clinical trial registration form were included. Although early warning systems are increasingly used in ED, compliance varies. One effectiveness study provided very low quality evidence (assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE)) that the use of an early warning system in the ED may lead to a change in patient management but does not reduce the number of adverse events; however this is uncertain, considering the quality of evidence. A total of 27 different early warning systems were developed/validated in 35 studies. There is relatively good evidence on the predictive ability of certain early warning systems on mortality and ICU/hospital admission. No health economic studies of health economic data in clinical studies were identified.

Conclusion

Early warning systems seem to be able to predict adverse outcomes in adult patients of varying acuity presenting to the ED, but there is a lack of high quality comparative studies to examine the effect of using early warning systems on patient outcomes. This should include a health economics assessment. Strategies for ensuring compliance should be developed and tested.

Table of contents

List of abbreviations

***Note the difference in the abbreviations for MEWS and I-MEWS**

+LR	Positive Likelihood Ratio
Aa gradient	Alveolar to arterial oxygen gradient
AAEM	American Academy of Emergency Medicine
ACDN	Alert and orientated, Confused, Drowsy, Not responsive or only to nail pressure
ACEM	Australasian College of Emergency Medicine
ACEP	American College of Emergency Physicians
ACT	Australian Capital Territory
AGREE II	Appraisal of Guidelines for Research & Evaluation
AHRQ	Agency for Healthcare Research and Quality
AMSTAR	Assessing the Methodological quality of Systematic Reviews
APACHE II	Acute Physiology and Chronic Health Evaluation score
ASEM	Australasian Society for Emergency Medicine
ASSIST	Assessment Score for Sick patient Identification and Step-up in Treatment
AUROC	Area Under the Receiver Operating Curve
AVPU	Alert, Voice, Pain, Unresponsive
BP	Blood Pressure
BEWS	Bispebjerg Early Warning Score
CBA	Controlled Before-and-After studies
CCI	Charlson comorbidity index
CD	Cannot be Determined
CEM	College of Emergency Medicine
CENNZ-NZNO	College of Emergency Nurses (New Zealand)
CI	Confidence Interval
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CURB-65	Confusion, Urea, Respiratory rate, Blood pressure, Age 65 or older
DIST	An Euclidean Distance-based Scoring System
EC	Emergency Call
ECG	Electrocardiogram
ED	Emergency Department
EDWIN	Emergency Department Work INdex
ED CIC	ED Critical Instability Criteria
EPOC	Effective Practice and Organisation of Care
ESI	Emergency severity index
ESS	Proposed Ensemble-Based Scoring System
eTTS	Electronically calculated Track & Trigger Score
EuSEM	European Society for Emergency Medicine
EWS	Early Warning Score
F_iO_2	Fraction of inspired oxygen
GCS	Glasgow Coma Scale
GIN	Guidelines International Network
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HDU	High Dependency Unit
HEED	Health Economic Evaluation Database
HIQA	Health Information and Quality Authority
HR	Heart Rate
HRV	Heart Rate Variability
HSE	Health Services Executive
IAEM	Irish Association for Emergency Medicine
ICER	Incremental Cost-Effectiveness Ratios (ICERs)
ICTRP	International Clinical Trials Registry Platform
ICU	Intensive Care Unit
IFEM	International Federation of Emergency Medicine

IHCA	In-Hospital Cardiac Arrest
IMEWS*	Irish Maternity Early Warning System
IQR	Interquartile Range
NEWS (Ireland)	Irish National Early Warning Score
ISBAR	Identify, Situation, Background, Assessment and Recommendation
ITS	Interrupted Time Series designs
LOC	Loss Of Consciousness
LODS	Logistic Organ Dysfunction System
MT	Medical Team
MEES	Mainz Emergency Evaluation Score
MeSH	Medical Subject Headings
MET	Medical Emergency Team
MEWS*	Modified Early Warning Score
MEWS plus	Modified Early Warming Score plus
MI	Myocardial Infarction
ML	Machine Learning
MEDS	Mortality in emergency department sepsis
MPM0 II	Morbidity Probability Model at admission
mREMS	Modified Rapid Emergency Medicine Score
MTS	Manchester Triage System
NCCHTA	National Coordinating Centre for Health Technology Assessment
NEDS	Nationwide Emergency Department Sample
NEWS	National Early Warning Score
NEWS-L	National Early Warning Score + Lactate
NHS	National Health Service
NHSEED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NIHR-HTA	National Institute for Health Research – Health Technology Assessment
NPT	Near-Patient-Test
NPV	Negative Predictive Value
NRCT	Non-Randomised Controlled Trial
OR	Odds Ratio
OTC	Over-The-Counter
PACS	Patient Acuity Category Scale
PARS	Patient At Risk Score
PEDS	Prince of Wales ED Score
PEWS	Paediatric Early Warning System
pH	Acidic/basic measure
PIRO	Predisposition, Insult/Infection, Response, and Organ dysfunction
POTTS	Physiological Observation Track and Trigger System
PPV	Positive Predictive Value
PSI	Patient Status Index
QALYs	Quality Adjusted Life Years
RAPS	Rapid Acute Physiology Score
RCEM	Royal College of Emergency Medicine
RCN	Royal College of Nursing
RCoP	Royal College Of Physicians
RCT	Randomised Controlled Trial
REMS	Rapid Emergency Medicine Score
ROB	Risk Of Bias
ROC	Receiver Operating Curve
RR	Risk Ratio
RTS	Revised Trauma Score
SAEM	Society for Academic Emergency Medicine

SAPS II	New Simplified Acute Physiology Score
SBP	Systolic Blood Pressure
SCS	Simple Clinical Score
SD	Standard Deviation
SIGN	Scottish Intercollegiate Guidelines Network
SIRS	Systemic Inflammatory Response Syndrome
SOFA	Sequential Organ Failure Assessment
SOS	Sepsis in Obstetrics Score
SS	Septic Shock
SSSS	Severe Sepsis and Septic Shock score
TC	Trauma Call
Temp	Temperature
TEWS	Triage Early Warning Score
THERM	The Resuscitation Management score
TIMI	Thrombolysis In Myocardial Infarction
TRISS	Trauma – Injury Severity Score
TTS	Track and Trigger System
UK	United Kingdom
VIEWS	VitalPAC Early Warning Score
VIEWS-L	VitalPAC Early Warning Score-Lactate
WBC	White Blood Cell count
WHO	World Health Organisation

List of tables

List of figures

1 Background

Serious clinical adverse events are related to physiological abnormalities and changes in physiological parameters such as blood pressure, pulse rate, temperature, respiratory rate, level of consciousness, often precede the deterioration of patients.[1-4] Early intervention may improve patient outcomes and failure to recognise acute deterioration in patients may lead to increased morbidity and mortality.[5, 6] Early warning systems and track and trigger systems (TTS) use routine physiological measurements, which are totalled to generate a score with pre-specified alert thresholds. Their aim is to identify patients at risk of deterioration early and trigger appropriate and timely responses, known as escalation of care.

Early warning systems are used increasingly in acute care settings and several countries have developed National Early Warning Scores (NEWS). In Ireland, the National Clinical Guideline on the use of NEWS for adult patients came into effect in 2013.[7] In the UK, The Royal College of Physicians (RCoP) published a National Early Warning Score in 2012,[8] and the National Institute for Health and Care Excellence (NICE) recommends the use of a TTS to monitor hospital patients.[9] In Australia, the Early Recognition of Deteriorating Patient Program introduced a TTS and corresponding educational programme.[10] Similarly, in the USA, Rapid Response Systems with fixed "Calling Criteria" are recommended to trigger adequate medical response.[11] Early warning systems have also been adapted to several specific contexts such as maternity care[12-14] and paediatric care; for example, the Irish Maternity (I-MEWS) and Paediatric Early Warning Systems (PEWS).[12, 15]

Many acutely ill patients first present to the emergency department (ED). The ED is a unique and complex environment, distinctly different from other hospital departments in many ways. Visits are unscheduled and patients attend with undiagnosed, undifferentiated conditions of varying acuity. Medical staff must care for several patients simultaneously, deal with constantly shifting priorities and respond to multiple demands due to the unpredictable nature of the ED environment.[16, 17] Initial triage determines the priority of patients' treatments based on the severity of their condition, but, following triage, continuous monitoring and prompt recognition of deteriorating patients is crucial to escalate care and transfer patients appropriately, particularly as there is a high prevalence of acute illness in the ED. Early warning systems are sometimes used as an adjunct to triage for early identification of deterioration in the ED, particularly in situations of crowding.[18] Common early warning systems such as the Modified Early Warning Score (MEWS)[19] are used frequently and validated against specific subgroups of patients (e.g. acute renal failure)[20] but may not be directly transferable to an ED setting[18] where patients present with a variety of unspecified conditions. There is, therefore, a need to review the use, effectiveness and cost-effectiveness of early warning systems specifically in an ED context to guide practice, future research and early warning system development.

2 Aim and objectives

The aim of this review is to provide a rapid synthesis[a] of the evidence of the clinical effectiveness and cost-effectiveness of physiologically based early warning systems and TTS for the detection of deterioration (post-triage) in adult patients presenting to ED.

The project addressed five main objectives:

1. To describe the use internationally, including the level of use and the variety of systems in use, of physiologically based early warning systems or TTS or scoring systems for the detection of deterioration in adult patients presenting to the ED;
2. To evaluate the clinical effectiveness of physiologically based early warning systems or TTS or scoring systems in adult patients presenting to the ED;
3. To describe the development and validation of such systems;
4. To evaluate the cost effectiveness, cost impact and resources involved in physiologically based early warning systems or TTS or scoring systems for the detection of deterioration in adult patients presenting to the ED;
5. To describe the education programmes, including the evaluation of such programmes that have been established to train healthcare professionals, and other non-professional staff, in the delivery of such systems.

a A rapid review has been defined as a systematic review that is limited in time and/or scope; however, there is no single definition of this capacity and large variation exists (31). This systematic review took place in a 12 week period with only minor restrictions to the methods used. The methods applied in this review are described fully in Section 3.

3 Methods

3.1 Selection criteria

3.1.1 Population, Intervention, Comparison, Outcome (PICO)

The PICO format was used to inform the search strategy according to the five objectives:

a . To describe the use internationally, including the level of use and the variety of systems in use, of physiologically based early warning systems or TTS or scoring systems for the detection of deterioration in adult patients presenting to Emergency Departments	
P	Adult patients presenting to the ED following initial triage. (Studies/reports that focussed on triaging patients or that were not set in the ED, were excluded.)
I	Early warning systems or TTS or scoring systems, relying on periodic observation of selected, routinely recorded, physiological parameters, to promptly recognise deteriorating patients and trigger escalation of care based on present response criteria.
C	N/A
O	Extent of use of early warning systems or TTS or scoring systems • Types of early warning systems or TTS or scoring systems in use • Number and type of clinical guidelines (regional, national, international)

b. To evaluate the clinical effectiveness of physiologically based early warning systems or TTS or scoring systems in adult patients presenting to the ED	
P	Adult patients presenting to the ED following initial triage.
I	Early warning systems or TTS or scoring systems, relying on periodic observation of selected, routinely recorded, physiological parameters, to promptly recognise deteriorating patients and trigger escalation of care based on pre-set response criteria.
C	Non-use of the systems or the use of alternative systems of physiological monitoring.
O	Clinical outcomes • Death • Critical illness (collapse – cardiac or respiratory arrest, haemorrhage, sepsis etc.) • Admission to intensive care unit (ICU) Length of hospital stay (days)

c. To describe the development and validation of such systems	
P	Adult patients presenting to the ED following initial triage.
I	Early warning systems or TTS or scoring systems, relying on periodic observation of selected, routinely recorded, physiological parameters, to promptly recognise deteriorating patients and trigger escalation of care based on pre-set response criteria.
C	N/A

O	Clinical outcomes
	• Death
	  • Critical illness (collapse – cardiac or respiratory arrest, haemorrhage, sepsis etc.)
	  • Admission to intensive care unit (ICU)
	  • Length of hospital stay (days)
	• Sensitivity of early warning systems or TTS or scoring systems for adverse outcome/critical illness criterion
	• Specificity of early warning systems or TTS or scoring systems for adverse outcome/critical illness criterion
	• Positive predictive value of early warning systems or TTS or scoring systems for adverse outcome/critical illness criterion
	• Negative predictive value of early warning systems or TTS or scoring systems for adverse outcome/critical illness criterion

d. To evaluate the cost effectiveness, cost impact and resources involved in physiologically based early warning systems or TTS or scoring systems for the detection of deterioration in adult patients presenting to the ED	
P	Adult patients presenting to the ED following initial triage.
I	Early warning systems or TTS or scoring systems, relying on periodic observation of selected, routinely recorded, physiological parameters, to promptly recognise deteriorating patients and trigger escalation of care based on pre-set response criteria.
C	Non-use of the systems or the use of alternative systems of physiological monitoring.
O	Economic measures of healthcare: • Use of healthcare resources associated with early warning systems or TTS or scoring systems use including direct medical resource costs (staff time, education time and cost, additional referrals), indirect costs (associated with loss of productivity) and other non-medical costs (e.g. patient out of pocket expenses) • Cost savings, cost effectiveness measures such as Incremental Cost-Effectiveness Ratios (ICERs), Quality Adjusted Life Years (QALYs).

e. To describe the education programmes, including their evaluation that have been established to train healthcare professionals, and other non-professional staff, in the delivery of such systems	
P	Healthcare professionals using physiologically based early warning systems or TTS or scoring systems and associated escalation protocols or communication tools in ED settings. Non-healthcare professional staff involved in the delivery of such systems.
I	Educational programmes for healthcare professionals concerning such early warning systems or TTS or scoring systems
C	Comparators included non-use or use of alternative educational programmes concerning early warning systems or TTS or scoring systems
O	• Types of education programmes • Strategies and methods to evaluate education programmes of early warning systems or TTS or scoring systems

3.1.2 Types of studies/reports

The following six types of studies or reports were included:

a. *Descriptive studies – types and use of systems:* Studies that described types or variety of early warning systems or TTS or scoring systems used and the extent to which they were used in clinical practice.

b. *Descriptive studies – education programmes:* Studies that described education programmes to train healthcare professionals in delivering early warning systems or TTS or scoring systems.

c. *Guidelines:* Regional, national and international guidelines that described early warning systems or TTS or scoring systems.

d. *Effectiveness studies:* Studies that examined the effectiveness of an early warning system or TTS or scoring system on outcomes for adults admitted to the ED following triage, and that had a controlled design (i.e., randomised controlled trials [RCTs], non-randomised controlled trials [NRCT], controlled before-and-after studies [CBA], interrupted time series designs [ITS] and cohort studies with historical controls). Studies that evaluated the effects of the system on relevant outcomes without control (e.g. case series, cohort studies without historical control), were included in the descriptive category.

e. *Development and validation studies:* Development studies were defined as studies that focused on the development of early warning systems or TTS or scoring systems while validation studies assessed the predictive ability of such systems. Studies in this category needed to include adult patients both with and without the reference outcome (such as admission to intensive care or mortality) or were otherwise considered a descriptive study. For the purpose of classification, we regarded studies as 'development' studies if reference ranges, parameters, and/or design of scoring systems were identified based on the outcomes of the study sample (for example, through the use of receiver operating characteristics [ROC] curves). In validation studies, such reference criteria were already determined and their predictive ability was evaluated in a new sample of patients.

f. *Health economics:* Full economic evaluation studies (cost-effectiveness analysis, cost-utility analysis and cost-benefit analysis), cost analysis and comparative resource use studies comparing early warning systems or TTS or scoring systems to one or more standard treatments. These may have included any study that met the eligibility criteria for the review of effectiveness; hence studies in other categories might have been also been included here.

3.2 Search methods

A comprehensive search was conducted for evidence on early warning systems, TTS or scoring system in ED, and included both database and grey literature searches. Individual search strategies were developed for four major electronic databases: the Cochrane Library (all databases therein), Ovid Medline, Embase and CINAHL. Additional resources that were searched included: specific cost-effectiveness resources (n=4), guidance resources (n=6), professional bodies resources (n=22), grey literature resources (n=3), and clinical trial registries (n=4). No language restrictions were applied, but considering this was a rapid review, no translation could take place. No filter was applied to the time of publication of resources. A filter (free text terms) was applied to limit retrieval to the adult population where available. Full details of individual search strategies, including the search dates, are provided in Appendix 1. Details of the search results are presented in a PRISMA flow diagram (Figure 1),[21] produced in RevMan.[22]

3.3 Screening for inclusion

Three reviewers (FW, PM and SD) screened the titles/abstracts from the database searches so that each citation was screened by at least two reviewers independently. For additional resources, the information specialist (AC) sifted the search results for potentially eligible studies (see Appendix 1). Full text reports from databases and from additional resources were assessed for inclusion, based on the selection criteria (section 3.1) by two reviewers independently (FW and PM) and discrepancies were resolved by discussion and, where necessary, by involving a third person (DD).

3.4 Risk of bias/methodological quality assessment

Two reviewers (from FW and/or VS and/or DD) independently assessed the Risk of Bias (ROB) and/or methodological quality of the included reports, using the critical appraisal instruments listed in Table 1.

Table 1. Risk of Bias and quality of evidence critical appraisal instruments

Study design	Risk of bias (ROB)/quality assessment tool
Descriptive studies	Adapted from National Institute of Health checklist[23]
Descriptive studies – educational programmes	Adapted from National Institute of Health checklist[23]
Guidelines	AGREE II tool[24]
Effectiveness studies – RCTs	Cochrane ROB tool[25] and GRADE quality of evidence assessment[26]
Effectiveness studies – non-RCTs	EPOC quality assessment for quantitative studies[25, 27] and GRADE quality of evidence assessment[26]
Systematic reviews	AMSTAR
Economic evaluations	British Medical Journal Checklist for authors and peer-reviewers of economic submission[28]; Checklist for quality assessment in economic decision-analytic models[29]
Development and validation studies	Quality Assessment Tool adapted from Kansagara et al (2011)[30]

3.5 Data extraction

Separate data extraction forms were designed for each of the six types of studies included in this review (section 3.1.4). Data extraction was completed by two reviewers (FW and PW). Each reviewer extracted data from half of the included reports and 50% of entries were checked by a second reviewer. The data elements that were extracted are presented in Table 2 below.

Table 2. Data extracted from included reports

Descriptive studies – types and use of systems
Authors, time and country of study Study aim and design Number of participants and characteristics Method(s) of data collection and analysis Content (parameters) of the early warning system or TTS or scoring system, and escalation criteria Findings on the use of early warning or track and trigger system(s)
Descriptive studies – education programmes
Authors, time and country of study Study aim and design Number of participants and characteristics Method(s) of data collection and analysis Content (parameters) of the early warning system or TTS or scoring system, and escalation criteria Information on the educational programme or communication tool Findings on the use of educational programme or communication tool concerning an early warning system or TTS or scoring system
Guidelines
Guideline team (including qualifications), time and country of guideline Guideline development strategy Scope Key recommendations Implementation strategy Audit strategy
Effectiveness studies
Authors, time and country of study Study aim and design Number of participants and characteristics Method(s) of data collection and analysis Intervention (content (parameters) of the early warning system or TTS or scoring system, and escalation criteria) and control Outcomes Findings, including effect estimates
Development and validation studies
Authors, time and country of study Study aim and design Number of participants and characteristics Method(s) of data collection and analysis Content (parameters) of the early warning system or TTS or scoring system, and escalation criteria Reference criteria (outcomes) Findings, including predictive ability measures

Health economics
Authors, time and country of study Study aim and design Number of participants and characteristics Method(s) of data collection and analysis Measures of cost Outcomes

3.6 Data analysis and synthesis

Data were collated in evidence tables for each of the six types of studies included in this review. In addition, we provide a concise narrative synthesis of the findings of descriptive studies, descriptive studies of educational programmes, guidelines, and development and validation studies related to early warning systems or TTS or scoring systems in ED. For effectiveness studies, a meta-analysis was planned but was not performed due to the limited number of studies (n=1), hence, a narrative summary is provided. For health economics studies, we planned to examine the cost-effectiveness of using early warning systems or TTS or scoring systems in ED, but no such studies were identified in the comprehensive search for this review.

3.7 Reporting of the review

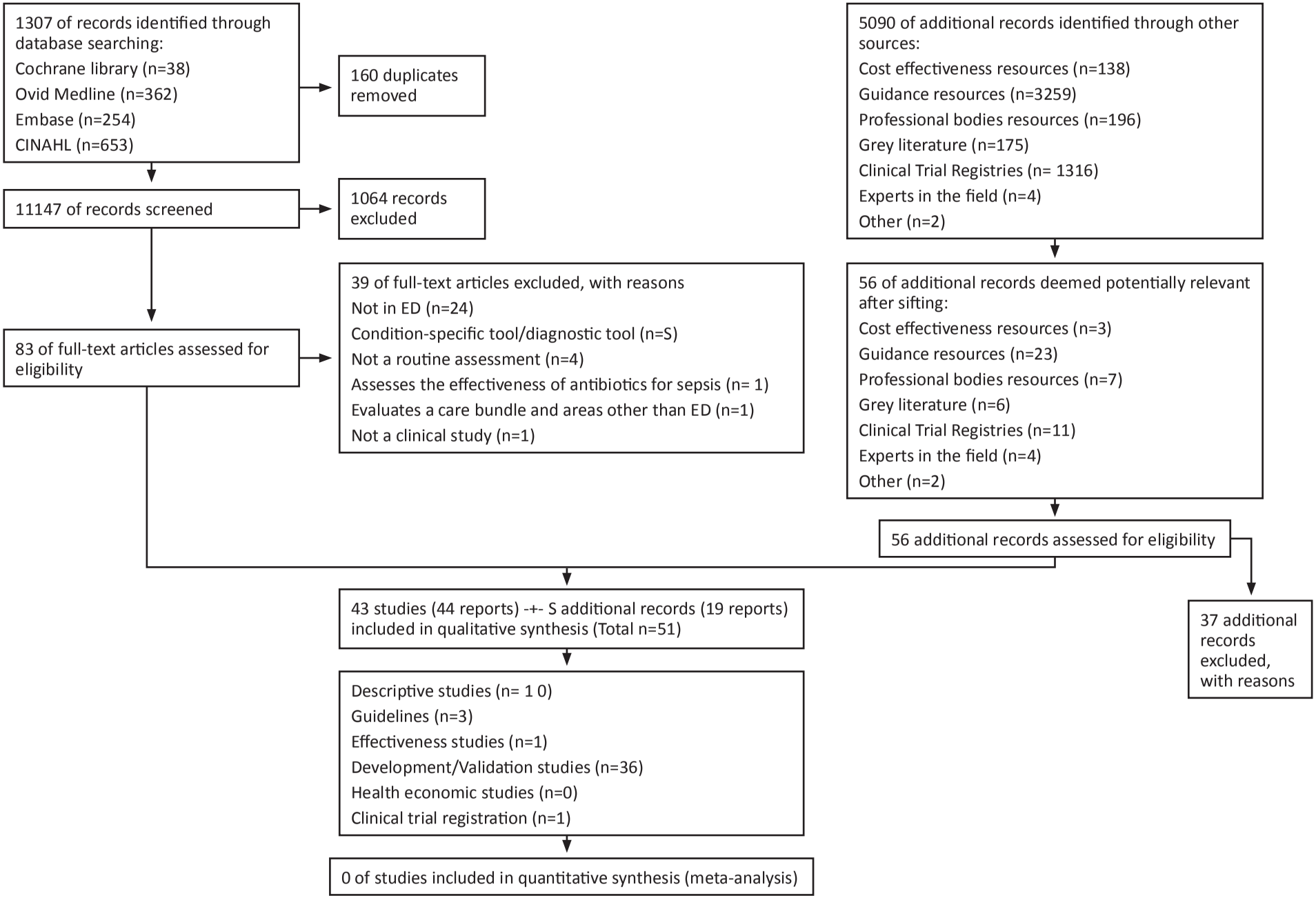
No reporting guidelines for rapid systematic reviews exist at present although one is currently in development (personal communication with D. Moher, Ottawa). The reporting checklist for brief reviews by Abrami et al[31] was completed to ensure adequate reporting (Appendix 2).

4 Results

4.1 Search results

A total of 6397 citations were identified (1307 from databases and 5090 from additional resources). After duplicate removal, 1147 database citations were screened against the selection criteria by title/abstract. Full texts of 83 reports were assessed of which 44 reports (43 studies) were finally included. The most common reason for exclusion was 'non ED setting' (n=24). One study in Chinese was identified and only data from the abstract (in English) was included.[32] Nineteen of the 56 screened additional resources were included, of these 14 reports related to three clinical guidelines and one to a registered trial. The results of the complete search strategy are presented in Figure 1.

Figure 1. Search Results



The number of included studies/reports by type of study/report (as specified in section 3.1) is presented in Table 3 below.

Table 3. Number and types of studies included in the review

Type of study/report	n
Descriptive studies – type, extent of use and compliance and use of systems (Note: 2 study sub-types (extent of use and compliance) emerged and are presented separately in tables 4 and 5 below)	10
Descriptive studies – educational programmes	0
Guidelines (and related documents)	3
Effectiveness studies	1
Validation & Development studies	35 (+1 review)
Health economics studies	0

4.2 Risk of bias and methodological quality of included reports

The risk of bias and quality assessment of the included reports is presented in Appendix 3. Details of the critical appraisal instruments used are outlined in table 1 above.

We examined the quality of six descriptive studies that examined the extent of early warning system(s) use. One report was a conference abstract and was rated of poor quality because only limited information was reported for quality assessment.[33] The remaining five studies were of fair quality.[34-37]

The four descriptive studies assessing compliance with using early warning systems were of good[38-40] and fair quality.[41]

One effectiveness study in the review was rated as high risk of bias overall. Shuk-Ngor et al [42] included a non-random sample and did not provide sufficient information to assess allocation bias and similarity in baseline characteristics between the two groups. Using the GRADE system for grading evidence, the quality of evidence of this study was very low for the effect of using MEWS compared to clinical judgement on a change in ED patient management and adverse events. This was because the study had a high risk of bias and considered a low number of few events.

Eight studies that developed and validated a system (in the same sample) were rated as having low (n=6) and unclear (n=2) risk of bias. The 27 studies that validated an existing system in a new cohort of people were judged as having low (n=16), unclear (n=8) and high risk of bias (n=3).[43-45] Studies with unclear risk of bias generally did not specify the methods of sampling (n=9), did not state cut-off values used (n=12) or did not pre-specify the outcomes clearly (n=1). One scoping review of predictive ability of early warning systems[46] was rated of good quality.

The three included guidelines were assessed using the AGREE II tool and scores given by the two reviewers independently were averaged for each domain. The Irish National Early Warning Score Guideline and the guideline of the National Institute for Health and Care Excellence were rated of high quality with the percentage scores for the six domains varying between 91.7 - 97.2% and 87.5 -100% respectively, and the overall quality score given by the reviewers was 91.7% for both guidelines.[7, 9] The guideline produced by the Royal College of Physicians was rated lower in quality (range 62.5-100%; overall quality 66.7%) mainly because of limited information available about their search strategy and

no clear statement of how recommendations were derived from evidence. They cite a systematic review as source of their evidence but the scope of the guideline was broader than that addressed in this review.[8]

4.3 Findings

The findings of this review are structured by type of study.

4.3.1 Descriptive studies – Type, extent of use, and compliance

Ten descriptive studies were included of which six examined the extent of using early warning systems[33-37, 47] and four examined compliance with such systems[38-41] One report was a conference abstract in which an early warning system was described but limited data were available.[33]

Extent of use

Six reports published in the last six years described the use of early warning systems within the ED (See Table 4 Evidence table below). The reports collected data from medical records,[34, 35] a survey,[36] a web-survey,[47] a prospective observational cohort study [37], and through participatory action research.[33] One report was a conference abstract in which the authors refer to a new monitoring system to identify the need for escalation of care, but the system was not described fully in the abstract.[33] Considine et al [34] described a pilot study in a hospital in Australia examining the use of an early warning system that considered criteria related to a patients' airway, circulation, disability and any sudden deterioration. The escalation protocol consisted of a review of the patient by an emergency physician within five minutes if any of the criteria were met, followed with additional interventions, if appropriate. A national survey in seven jurisdictions in Australia, found that 20 of 220 hospitals had a formal rapid response system in the ED, but the prevalence of early warning systems in EDs was not reported.[47] We did not find any other studies reporting any aspects of the escalation protocols.

Wilson et al [37] included the parameters heart rate (HR), blood pressure, respiratory rate, peripheral oxygen saturation, temperature and the Glasgow Coma Scale (GCS) in their TTS hospital chart. They compared the TTS scores recorded in the charts with scores calculated retrospectively and found that 20.6% (n=211) were incorrect. This was mainly because of incorrect assignment of the score to an individual vital sign, which led to underscoring of the total TTS and reduced escalation activation. Correia et al [35] did not provide details on the content of the early warning system they used in a small study (n=69) in Portugal, but found that a threshold of a score ≥ 3 would have increased early medical attention by 40% compared to clinical judgement alone. A survey in 2012 of 145 (57% response rate) clinical leads of EDs in the UK showed that 71% used an early warning system, with the Modified Early Warning Score (MEWS) being the most common system used (80%).

In summary, multiple early warning systems are available and the extent of their use in the ED may vary geographically but limited data precludes comparisons between countries. Some evidence suggests that incorrect calculation or recording of scores may lead to inappropriate escalation activation or a lack of such activation, drawing attention to the importance of adequate implementation of such systems.

Compliance

Three retrospective studies[38-40] and one audit[41] conducted in the UK, Denmark, the USA and New Zealand in the last five years examined compliance with recording early warning system parameters and escalation of care (See Table 5 Evidence table below). The vital sign parameters included in the early warning system were respiratory rate, HR, systolic blood pressure (SBP), temperature and level of consciousness (LOC) for one study by Christensen et al [38]. Another study by Hudson et al [41] also included urinary output, pain scores and the presence of recurrent/prolonged seizures or uncontrollable/

new pain, in addition to the vital sign parameters. Austen et al [40] included urine output and oxygen saturations as well as the vital sign parameters. Christensen et al [38] reported a rate of 7% (22/300) of calculated scores in the clinical notes; however, 16% of records included all five vital signs, and HR, SBP, and LOC were reported in 90-95% of records. Compliance with escalation of care varied; all nine patients that met the trauma call activation criteria (immediately life-threatening signs/symptoms or BEWS ≥5) had triggered a trauma call, but only 24 of the 48 emergency call activation criteria (immediately life-threatening signs/symptoms or BEWS ≥5) had been responded to by an emergency call. Austen et al [40] found a much higher compliance with 66% of records containing an aggregate score, although only 72.6% of these were accurate. In an audit, the pre-implementation rate (30%) of abnormal vital sign identification was significantly lower than the post-implementation (53.5%) rate (p=0.007) and patients were less likely to receive medication (p=0.001), but no details of the implementation strategy they used were described.[41]

In summary, four studies examined compliance and the factors affecting monitoring vital signs in an ED setting. Compliance with recording and responding to early warning systems appeared relatively low although this varied across the included studies. The rate of vital sign monitoring for some but not all individual vital signs was high. The frequency of recording of HR and BP were particularly high, but the frequency of recording of temperature (65.0%-96.8%) and respiratory rate (18.0%-98.9%) varied across the included studies. While certain factors, including patients' triage category, age, and number of medications seem to increase frequency of vital sign monitoring, it also appears that crowding at the ED, increased length of time in the ED and a decreased number of routes of medication administration may lead to reduced monitoring.[39]

Table 4. Evidence Table: Descriptive studies - Extent of use

(Methodological quality was rated using an adapted National Institute of Health checklist.[23] Full details of the methodological assessment are available in Appendix 3.)

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/ tool	Results
Australian Commission on Safety and Quality in Health Care (2011),[47] Australia **Quality Rating:** Fair	To describe recognition and response systems in Australian hospitals.	Web-based Survey	Public and private hospitals in 7 Australian jurisdictions between Sept and Dec 2010. 227 nominees were provided to the Commission, and 182 (representing 220 hospitals; 143 public and 77 private) of these (80%) completed the survey.	Examines extent of use of systems. (No details provided on individual systems.)	**Systems for recognising clinical deterioration** • 77% had written policies, protocol or guidelines regarding the measurement of physiological observations (% for ED not stated) • 77% had a formal escalation protocol (of which 45% had a graded response) • 35% used formal early warning system or TTS (of which 58% were single or multiple parameter systems, 10% a TTS that required score calculation, and 26% used a combined system). • 50% used a structured protocol or tool for handover communications (common tools were SBAR (34%) ISOBAR (33%) and ISBAR (21%)). **Systems for responding to deterioration** • 66% had a formal rapid response system (larger hospitals and those in metropolitan areas were more likely to have these systems), of which 24% (n=20) were in ED • In 100% of hospitals, nurses on the ward could call the rapid response system. Doctors on call in could call the rapid response system in 89% of hospitals, other hospital staff in 69%, and families, patients and carers in 18% of hospitals. **Organisational systems to support the recognition of and response to deterioration** • 70% had identified staff in their hospitals with primary responsibility for developing, implementing, sustaining and monitoring recognition and response systems • 6% had been allocated specific funding for the operation of their rapid response system

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/ tool	Results
					• 72% had a committee that oversaw the operation of these systems • 69% provided regular training and education to support staff • 48% collected specific data about the effectiveness of their recognition and response systems
Considine et al (2012),[34] Australia **Quality Rating**: Fair	Evaluate the uptake of ED Early Warning Score for recognition of and response to clinical deterioration.	Pilot Descriptive exploratory study	300-bed urban district hospital. Systematic sample of 204 patients for whom ED EWS had been activated (every 10th patient in ED EWS log book over 24 months period).	**Critical instability criteria** • Airway/breathing: Stridor, upper airway obstruction, or threatened airway, SpO2 < 90%, Arterial blood gases pH < 7.20, Respiratory rate < 10 or > 30 breaths/ min • Circulation: Heart rate < 50 or > 120 beats/min, SBP < 90 or > 200 mmHg, Urine output < 20 or < 100 mL/6 h • Disability: Sudden decrease in consciousness (fall in GCS score > 2), Repeated or prolonged seizures • Worried?: Patients who may not meet above criteria but have a sudden deterioration, requiring urgent medical review.	• 204 patients (of which 16 patients <16 years) • Nurses made 93.1% of ED EWS activations. • Most common reasons for ED EWS activation were: respiratory (25%) and cardiac (22.5%) (Hypotension (27.7%) and tachycardia (23.7%) were most common reasons for ED EWS activation.) • 82.4% of patients were seen by medical staff before ED EWS activation. • Median duration of clinical instability was 39 minutes (IQR, 5 – 129 minutes). • Median time between documenting physiological abnormalities and ED EWS activation was 5 minutes (range 0 – 20). • Most patients (57.8%) required hospital admission: 4.4% of patients required ICU admission.

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/ tool	Results
Correia et al (2014),[35] Portugal **Quality Rating**: Fair	To assess the Early Warning Score (EWS) in specific time windows preceding an acute event, to study its temporal behaviour and its relation to outcomes, to compare it with established ward care.	Retrospective cohort	First consecutive 100 adult ward patients assisted by the outreach team and transferred to ED from 1 Jan to 31 April 2009.	EWS: parameters not clearly specified. Score threshold of >3 as trigger.	n=65 (65% of eligible sample) • Main cause of deterioration: Respiratory problems (44.6%); cardiovascular (27.7%) and neurological deterioration (27.7%). • EWS score at three periods preceding ward transfer to the ED (EWS Mean/SD): 72 h: 2.6 ± 1.9 24 h: 2.4 ± 1.8 12 h: 3.8 ± 1.7 • Score at 24h and 12h seemed to predict both length of stay and mortality ($p < 0.05$). • 63% were admitted in ICU or Intermediate Care Units* (26% and 37%, respectively), 20% returned to their origin wards, and 17% died in the ED. The overall in-hospital mortality = 53.8% • The EWS would have increased early medical attention by 40% if a threshold of ≥3 was used. * This study describes 3 levels of care (return to ward, admission to Intensive or Intermediate Units)
Coughlan et al (2015),[33] Ireland (Conference abstract) **Quality Rating**: Poor	To provide a novel longitudinal monitoring system tailored to identify an escalation protocol in ED	Participatory Action Research	NR	Describes a new system (see results)	Monitoring and response system includes: a) Monitoring chart for adult patients b) Standardised approach to monitoring & reassessment of patients after triage until medical assessment c) ISBAR tool d) Template for patient specific monitoring plan e) Template for escalation protocol

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/ tool	Results
Griffiths et al (2012),[36] UK **Quality Rating**: Fair	To assess the use of early warning systems in UK EDs and whether the respondent supported the use of early warning systems in the ED.	Survey	254 adult ED clinical leads	Examines extent of use of systems. (No details provided on individual systems.)	Response rate = 57% (145/254). Of the 145, 87% used an early warning system. • 71% used early warning systems to trigger senior review. • Types of early warning systems used: 80% MEWS, 10% PARS, 10% other (MEWS & PARS are both aggregate scores) • In 76% patients with high scores are cared for with increased monitoring. • In 44% increased scores triggered critical care input. • 93% of clinical leads supported the use of early warning systems in ED. • Discharge of patients with high scores: consultant review (32%), admitted (22%), no senior review (23%), other (17%), no answer (6%).
Wilson et al (2013),[37] UK **Quality Rating**: Fair	To evaluate the utilisation of paper based Track and Trigger (TTS) charts in a UK emergency department.	Prospective observational cohort	472 adults (over 18 years) entering one of three clinical areas of the ED (resuscitation room, 'majors',observation ward)	**Vital sign and TTS data:** Parameters: Heart rate (HR), systolic and diastolic blood pressure (BP), respiratory rate, peripheral oxygen saturation (SpO_2), temperature and Glasgow Coma Scale (GCS) score.	**Completion of observations** 85.8% had ≥1 set of observations documented to the College of Emergency Medicine (CEM) standard of six parameters (HR, respiratory rate, BP and SpO_2, temperature, GCS). **Completion of TTS scores** 60.6% had ≥1 TTS score documented in ED 34.5% of observations contained a TTS score, of which 20.6% (211) were incorrect (79.1% of the incorrect TTS totals were underscored, potentially preventing a trigger event from being recognised; 93.4% of the errors can be solely attributed to the incorrect assignment of the score to an individual vital sign; incorrect addition of individual TTS scores occurred in 2.8% of errors). **Escalations** ≥1 escalation: 204 (escalation at ED arrival (n=163 with red/orange triage), of which 37 had 2nd escalation; escalation after arrival (n=41), of which 9 had 2nd escalation).

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/ tool	Results
					Completion of paper TTS charts ('Real TTS') TTS score exceeding alerting threshold: Escalation (n=29); No escalation (n=22) TTS score not exceeding alerting threshold: Escalation (n=94); No escalation (n=141) TTS scores not calculated: Escalation (n=81); No escalation (n=105) **Retrospective TTS completion ('Potential TTS')** TTS score exceeding alerting threshold: Escalation (n=110); No escalation (n=80) TTS score not exceeding alerting threshold: Escalation (n=94); No escalation (n=188)

Table 5. Evidence table: Descriptive studies – Compliance

(Methodological quality was rated using an adapted National Institute of Health checklist 23. Full details of the methodological assessment are available in Appendix 3.)

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/tool	Results
Austen et al (2012), UK **Quality Rating**: Good	To assess the degree of adherence to the Chelsea Early Warning Score (CEWS).	Retrospective chart review	94 conveniently sampled patient records (only ED data included in this review).	**CEWS** Parameters: Respiratory rate, oxygen saturations, temperature, SBP, HR, LOC (AVPU), urine output	**% of patients in whom parameter was recorded**: • Temperature: 96.8% • HR: 100% • Blood pressure: 100% • Respiratory rate: 98.9% • Oxygen saturations: 97.9% • AVPU: 96.8% • Urine output: 47.9% **% of patient with aggregate CEWS recorded**: 66.0% (62/94) **% of patient with aggregate CEWS correctly calculated**: 72.6% (45/62) Resulted in 10 patients for whom care was not escalated but should have.
Christensen et al (2011),[38] Denmark **Quality Rating**: Good	Examine whether the Bispebjerg EWS (BEWS) triage system is used systematically and correctly in a mixed ED population.	Retrospective cross sectional analysis	600-bed urban teaching hospital. 300 randomly selected 'red' (most severely ill/injured) category patients over a 6 month period in 2009 (=1/9 of total 'red' population during study period).	BEWS: Score 0-3 on 5 vital signs: Resp rate, HR, SBP, Temp, LOC. BEWS ≥ 5 activated emergency or trauma call	• BEWS calculated in notes: 7% (n=22/300) • HR, BP, LOC documented in 90-95% of cases; temp in 65% of cases; Respiratory rate in 18% of cases. • All 5 vital signs documented in 16.0% of cases. • Trauma Call (TC) activation criterion (n=9); in all these cases a TC was activated. • Emergency call (EC) EC activation criterion (n=48), but an EC was only activated in 24 patients. Among the 24 patients for whom an EC had not been activated, eight had a "primary criterion" (life-threatening signs/symptoms) and 16 patients had a retrospective BEWS ≥ 5.

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/tool	Results
Hudson et al (2015),[41] New Zealand **Quality Rating**: Fair	To standardise an emergency observation chart, the Adult Emergency Department Flow Chart, which incorporates elements designed to allow clinicians to more readily recognise the trends of patient deterioration.	Audit	181 medical records, randomly selected from two metropolitan hospitals: • 80 during the pre-implementation audit • 101 during the post-implementation audit.	Adult Emergency Department Flow Chart: Parameters: SBP, pulse rate, respiratory rate, urinary output, pain score, new/change/ uncontrollable pain, Glasgow Coma Score, Recurrent/Prolonged Seizures, O_2 saturation	• Patients at the pre-implementation audit were less likely to be identified as having an abnormal vital sign when compared to those at the post-implementation audit (30% pre, 53.5% post; Chi square = 14.261, p = 0.007). • Time taken from triage to identification of individual abnormal vital signs: no differences (p-values range from 0.2 to 0.5). • Time taken from identification of an abnormal vital sign to its subsequent management: slightly longer at the pre-implementation audit (Mean = 40 minutes, SD = 57 minutes) than at the post-implementation audit (Mean = 30 minutes, SD = 44 minutes), but no statistically significant difference. • Documentation of a medical officer being in attendance: Lower at the pre-implementation audit (Mean = 0.2, SD = 0.5) than at the post-implementation audit (Mean = 0.5, SD = 0.7; t = 2.6, p = 0.01). • Patients at the pre-implementation audit were less likely to receive medication to manage their abnormal vital sign (Pre: Mean = 0.2, SD = 0.4; Post: Mean = 0.4, SD = 0.6; t = 3.3, p = 0.001).

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/tool	Results
Johnson et al (2014),[39] USA **Quality Rating**: Good	(1) What are the personal health factors (number of prescription medications, number of OTC medications, comorbidities, age, gender, triage category) that affect the frequency of vital sign monitoring in the emergency department? (2) What social factors (insurance status, ethnicity) affect the frequency of vital sign monitoring in the emergency department? (3) Does the effect of personal factors on the frequency of vital sign monitoring in the emergency department change when environmental factors (family presence, crowding level, length of stay, number of routes of medications administered in emergency department) are taken into account?	Descriptive, retrospective chart review	Selected 165 charts from a possible 3,727 subjects from the crowded periods Emergency Department Work Index (EDWIN ≥2) and 60 of a possible 73 subjects from non-crowded periods (EDWIN <2), for a total of 225 reviewed charts.	**Assessment of impact of following factors on vital sign monitoring frequency:** <u>Personal health factors</u> No. of prescription medications No. of OTC medications No. of comorbidities Age Gender Triage category <u>Social factors</u> Ethnicity Insurance <u>Environmental factors</u> EDWIN (measure of crowding) Length of stay Family presence Routes of medication	**Influence of personal health factors** • Number of prescription medications ($p<0.01$), comorbidities, age ($p<0.01$), gender ($p<0.05$), and triage category ($p<0.001$) had significant correlation with the frequency of vital sign monitoring. • Strongest predictor of the frequency of vital sign monitoring: triage category ($t = 2.1$, $P = 0.04$). • Triage category had the greatest impact on the time between vital signs. (For every increase of 1 in the triage category (becoming less acute), the time between vital signs was increased by 34 minutes.) **Influence of social factors** None correlated with the frequency of vital sign monitoring. **Influence of environmental factors** Crowding level ($t = 2.3$, $P = 0.02$), length of stay ($t = 2.7$, $P = 0.008$), and number of routes of medications ($t = –2.5$, $P = 0.02$) were found to be significant predicators. As the EDWIN score increased by 1, the length of time between recording vital signs increased by 1.5 minutes.

4.3.2 Descriptive studies – Educational programmes

We did not identify any studies that described educational programmes related to early warning systems or TTS, although the three guidelines included in the review contain an educational tool (see section 4.2.3).

4.3.3 Guidelines

Three clinical guidelines were identified; one from Ireland [7] and two from the UK,[8, 9] published between 2007 and 2013 (See Table 6 Evidence table below). None were specific to an ED setting; all three apply to all acutely ill adult patients, but do not specifically exclude ED. The Irish National Early Warning Score (Irish NCG No. 1 NEWS)[7] and the NEWS guideline of the UK Royal College of Physicians (RCoP)[8] both include appended observation charts using a colour system to trigger escalation of care when appropriate, whereas the UK National Institute for Health and Care Excellence guideline (NICE CG50) recommend the use of a TTS but do not provide an exemplar chart.[9]

All three guidelines recommend the same six parameters to be measured: respiratory rate, heart rate, systolic blood pressure, temperature, oxygen saturations, and level of consciousness. The Irish NCG No. 1 NEWS also takes into account if a patient is on inspired oxygen (F_iO_2) and the NICE guideline includes a statement on additional parameters such as urine output in certain circumstances.

Each guideline is accompanied by an educational tool. The Irish NCG No. 1 NEWS guideline adapted the COMPASS educational program (Health Directorate ACT Government, Australia) for Irish use, and the NICE and RCoP guidelines developed online learning tools. Only the Irish NCG No. 1 NEWS and NICE guidelines provide a clear audit strategy for implementation of the guideline.

In summary, current guidelines on early warning systems for monitoring acute patients are not specific to the ED context. Common parameters across existing guidelines include respiratory rate, heart rate, systolic blood pressure, temperature, oxygen saturations, and level of consciousness.

Table 6. Evidence table: Guidelines

(Methodological quality was rated using the AGREE II tool.[24] Full details of the methodological assessment are available in Appendix 3.)

Authors (year), country	Population/ setting	Content of system	Frequency of recording	Escalation protocol	Implementation strategy	Audit strategy	Educational tool	Summary of key recommendations from guideline
Department of Health, National Clinical Guideline No.1 NEWS (2013),[7] Ireland **AGREE II score**: 91.7%	All adult patients in acute hospitals (excludes obstetric patients).	Respiratory rate Oxygen saturation (SpO_2), Heart rate, Blood pressure, Temperature, Level of consciousness. Where a patient is on inspired oxygen (F_iO_2) a score of 3 is added.	Yes (Minimum observation frequency is provided by score).	Yes (Recommendations are made but it is the responsibility of each individual hospital to outline their escalation protocol).	Yes (roles/ response-abilities and barriers/ enablers identified)	Yes (Tool provided)	Yes (COMPASS educational tool (online manual for independent learning + multiple choice quiz + face-to-face session), and ISBAR (Identify, situation, background, assessment, recommendations communication tool)	There are 60 recommendations; no key recommendations listed. (Full guideline available at http://health.gov.ie/wp-content/uploads/2015/01/NEWSFull-ReportAugust2014.pdf)
Royal College of Physicians; NEWS; report of a working party,[8] UK **AGREE II score**: 66.7%	Acutely ill adult patients in hospital, but also prehospital	1 respiratory rate 2 oxygen saturations 3 temperature 4 systolic blood pressure 5 pulse rate 6 level of consciousness.	Yes (Minimum observation frequency is provided by score).	Yes (Clinical responses are provided for the different scores)	Yes (recommend training should be mandatory for healthcare professionals and students, and then refer to educational tool)	No	Yes (e-learning sessions & field specific case studies)	NEWS should be used when patients present acutely to hospital and in pre-hospital assessment i.e. by primary care and ambulance services. The new National Early Warning Score (NEWS) report, which advocates standardising the use of a NEWS system across the NHS to drive the 'step change' required in the assessment and response to acute illness. NEWS could also be adopted as a surveillance system for all patients in hospitals for tracking their condition, alerting the clinical team to medical deterioration and triggering a timely response. (Full guideline at https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news)

Authors (year), country	Population/setting	Content of system	Frequency of recording	Escalation protocol	Implementation strategy	Audit strategy	Educational tool
National Institute for Health and Care Excellence CG50,[9] UK **AGREE II score**: 91.7%	Acutely ill adult patients in hospital.	Multiple-parameter or aggregate weighted scoring systems used for track and trigger systems should measure: • heart rate • respiratory rate • systolic blood pressure • level of consciousness • oxygen saturation • temperature. In specific clinical circumstances, additional monitoring should be considered; for example: • hourly urine output • biochemical analysis, such as lactate, blood glucose, base deficit, arterial pH • pain assessment.	No specific chart provided. Recommended to use a track and trigger system.	No specific service configuration can be recommended as a preferred response strategy for individuals identified as having a deteriorating clinical condition. A graded response strategy for patients identified as being at risk of clinical deterioration should be agreed and delivered locally. It should consist of the following three levels: low, medium and high score groups.	Yes (Implementation tools provided: https://www.nice.org.uk/guidance/cg50/resources/implementation-advice-433575469)	Yes (Tool provided)	Yes (presenter slides and shared e-learning program)

Summary of key recommendations from the NICE guideline (Full guideline available at https://www.nice.org.uk/guidance/cg50)

1. Adult patients in acute hospital settings, including patients in the emergency department for whom a clinical decision to admit has been made, should have:
   a. physiological observations recorded at the time of their admission or initial assessment
   b. a clear written monitoring plan that specifies which physiological observations should be recorded and how often. The plan should take account of the:
      - patient's diagnosis
      - presence of comorbidities
      - agreed treatment plan

Physiological observations should be recorded and acted upon by staff who have been trained to undertake these procedures and understand their clinical relevance.

2. Physiological track and trigger systems should be used to monitor all adult patients in acute hospital settings.
   a. Physiological observations should be monitored at least every 12 hours, unless a decision has been made at a senior level to increase or decrease this frequency for an individual patient.
   b. The frequency of monitoring should increase if abnormal physiology is detected, as outlined in the recommendation on graded response strategy

3. Staff caring for patients in acute hospital settings should have competencies in monitoring, measurement, interpretation and prompt response to the acutely ill patient appropriate to the level of care they are providing. Education and training should be provided to ensure staff have these competencies, and they should be assessed to ensure they can demonstrate them.

4. A graded response strategy for patients identified as being at risk of clinical deterioration should be agreed and delivered locally. It should consist of the following three levels.
   a. Low-score group:
      - Increased frequency of observations and the nurse in charge alerted.
   b. Medium-score group:
      - Urgent call to team with primary medical responsibility for the patient.
      - Simultaneous call to personnel with core competencies for acute illness. These competencies can be delivered by a variety of models at a local level, such as a critical care outreach team, a hospital-at-night team or a specialist trainee in an acute medical or surgical specialty.
   c. High-score group:
      - Emergency call to team with critical care competencies and diagnostic skills. The team should include a medical practitioner skilled in the assessment of the critically ill patient, who possesses advanced airway management and resuscitation skills. There should be an immediate response.

5. If the team caring for the patient considers that admission to a critical care area is clinically indicated, then the decision to admit should involve both the consultant caring for the patient on the ward and the consultant in critical care.

6. After the decision to transfer a patient from a critical care area to the general ward has been made, he or she should be transferred as early as possible during the day. Transfer from critical care areas to the general ward between 22.00 and 07.00 should be avoided whenever possible, and should be documented as an adverse incident if it occurs.

7. The critical care area transferring team and the receiving ward team should take shared responsibility for the care of the patient being transferred. They should jointly ensure:
   a. there is continuity of care through a formal structured handover of care from critical care area staff to ward staff (including both medical and nursing staff), supported by a written plan
   b. that the receiving ward, with support from critical care if required, can deliver the agreed plan.

The formal structured handover of care should include:
   c. a summary of critical care stay, including diagnosis and treatment
   d. a monitoring and investigation plan
   e. a plan for on-going treatment, including drugs and therapies, nutrition plan, infection status and any agreed limitations of treatment
   f. physical and rehabilitation needs
   g. psychological and emotional needs
   h. specific communication or language needs.

4.3.4 Effectiveness studies

Only one study examined the effectiveness of early warning systems and TTS (see Table 7 Evidence table below). A non-randomised controlled design compared the effect of the Modified Early Warning Score (MEWS) with clinical judgment on changes in the management and adverse events of patients who are waiting for in-patient beds in ED of a large hospital in Hong Kong.[42] The authors concluded that the MEWS might improve the rate of activating a critical pathway but might make little or no difference to the detection of deterioration or adverse events; however, we are very uncertain since the evidence was of very low quality (GRADE) due to serious imprecision and high risk of bias (Appendix 3).

In summary, there is limited evidence regarding the effectiveness of using early warning systems in the ED and available evidence from one study is of very low quality, making conclusions uncertain.

Table 7. Evidence Table: Effectiveness studies

(Risk of bias and quality of evidence were rated using the EPOC quality assessment for quantitative studies[27] and GRADE.[26] Full details of the assessment are available in Appendix 3.)

Authors (year), country	Study aim	Participants	Intervention	Control	Outcomes	Results
Shuk-Ngor et al (2015),[42] Hong Kong **Design**: non-randomised controlled trial **Risk of bias:** High **GRADE level**: Very low	To compare the performances of detecting patient deterioration with and without using the Modified Early Warning Score (MEWS) for a group of patients who are waiting for in-patient beds in an ED.	Emergency patients being held in the ED observation area because of access block to the specialty wards were included in the study. 544 patients recruited: • Intervention group (MEWS): 269 patients • Control (Usual Observation): 275 patients	Emergency nurses recorded the MEWS of access block patients 4-hourly, and followed MEWS action pathway for decision making to trigger actions: MEWS = 0—3 Regular observation MEWS = 4 Senior emergency nurse reviews patient's condition MEWS > 4 Triggered the critical pathway. Any time a nurse found that a patient was unwell, he/she could override a negative MEWS decision (i.e. MEWS < 5) to trigger the critical pathway.	''Clinical judgment'' = practice by nurses using individual's knowledge, clinical experience and gut feeling — judging based on strong feelings rather than facts, plus the measurement of 3 vital signs — blood pressure, pulse and body temperature.	**Primary outcome:** A change in patient's ED management plan by ED doctor in response to the MEWS critical pathway activation. **Secondary outcome:** An adverse event occurred during the first 24 h of admission to the ward. Adverse events were defined as active resuscitation, ICU admission, cardiac arrest and death.	**Change in management:** A positive doctor response = following the critical pathway activation, a senior doctor reviewed a patient and changed the management plan of that patient. A negative doctor response = no change in ED management plan following critical pathway activation. In the MEWS group, there was approximately 1 episode of activation in every 10 patients but it was 1 in 20 patients in the Usual Observation group. The positive doctor response rate was high in both patient groups (87.1% in the MEWS group; 92.9% in the Usual Observation group). **Adverse events:** 0.4% (n=1) in the MEWS group and 0.4% (n=1) in the Usual Observation group, had an adverse event within 24 h of admission to the ward. **Observation performance (Detection patient deterioration):** <u>MEWS:</u> Sensitivity: 100.0% Specificity: 98.3% <u>Usual observation:</u> Sensitivity: 100.0% Specificity: 97.8%

4.3.5 Development and validation studies

Challen and Goodacre [46] reported the results of a scoping review (See Table 9 evidence table below), which identified 119 tools related to outcome prediction in ED; however, the majority were condition-specific tools (n=94). They found the APACHE II score to have the highest reported area under the receiver operating characteristic (AUROC) curve[b] (0.984) in patients with peritonitis.

In addition, 35 development and/or validation primary study reports were identified (see Table 10, Table 11 and Table 12 Evidence tables below). Studies were conducted between 2003 and 2016 in the UK (n=5),[43, 49-52] the USA (n=5),[53-57] Turkey (n=4),[58-60] Hong Kong (n=3),[61-63] Singapore (n=3),[64-66] South Africa (n=2),[44, 45] Sweden (n=2),[67, 68] Denmark (n=1),[69] Germany (n=2),[70, 71] China (n=1),[32] South Korea (n=2),[72, 73] Thailand (n=1),[74] Taiwan (n=1),[75] the Netherlands (n=1)[76] and Australia (n=2).[77, 78] Twelve studies were retrospective, 22 were prospective cohort studies and one was a secondary analysis of a RCT.[56] Eight studies developed and validated (in the same sample) an early warning system, while 27 validated an existing system in a different sample. All 35 studies examined the use of early warning systems in an ED population. Three studies included a random sample[54, 69, 78] and participants in the remaining studies were recruited consecutively or the sampling strategy was not stated clearly.

A total of 27 early warning systems were developed and/or validated. Condition-specific systems; for example, the Mortality in Emergency Department Sepsis (MEDS) score, Sepsis in Obstetrics Score (SOS), CURB-65, and Trauma and Injury Severity Score (TRISS) were excluded from this review. Tools included were: the Modified Early Warning Score (MEWS),[32, 45, 51-54, 58-66, 70, 71, 74, 75, 79] the Rapid Emergency Medicine Score (REMS),[53, 59, 61, 62, 67, 68] the Prince of Wales ED Score (PEDS),[61, 62] the Revised Trauma Score (RTS),[61] the Acute Physiology and Chronic Health Evaluation score (APACHE II),[57, 61, 67, 77] The Resuscitation Management score (THERM),[62] the Simple Clinical Score (SCS),[62] the Mainz Emergency Evaluation Score (MEES),[62] National Early Warning Score (NEWS),[49, 50, 62, 72, 76] the Bispebjerg EWS (BEWS),[69] the Charlson comorbidity index (CCI),[60, 71, 75] the Emergency severity index (ESI),[71] MEWS plus,[54] modified REMS (mREMS),[55] National Early Warning Score including Lactate (NEWS-L),[72] the New Simplified Acute Physiology Score (SAPS) II,[56, 77] the Morbidity Probability Model at admission (MPM0 II),[56] the Logistic Organ Dysfunction System (LODS),[56] the Triage Early Warning Score (TEWS),[44] the Predisposition, Insult/Infection, Response, and Organ dysfunction (PIRO) model,[57] the Rapid Acute Physiology Score (RAPS),[67, 68] the Assessment Score for Sick patient Identification and Step-up in Treatment (ASSIST),[51] the Sequential Organ Failure Assessment (SOFA),[77] the Patient Status Index (PSI),[43] the VitalPAC Early Warning Score (VIEWS),[79] the VitalPAC Early Warning Score-Lactate (VIEWS-L)[73] and the ED Critical Instability Criteria (ED CIC).[78]

Churpek et al [80] classified early warning systems into single-parameter systems, multiple-parameter systems and aggregate weighted scores. A single-parameter system consists of a list of individual physiologic criteria that, if reached by a patient, triggers a response. Multi-parameter systems use combinations of physiologic criteria without calculation of a score to activate a response, while aggregate systems categorise vital signs and sometimes other variables into different degrees of physiologic abnormality and then assign point values for each category. The early warning systems examined in the studies included in this review primarily developed/validated aggregate weighted scores (Table 8).

b The Receiver Operating Curve (ROC) plots the true positive rate against the false positive rate at certain thresholds. The AUROC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance.48. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters. 2006;27: 861–74.

Table 8. Early warning systems included in the review by type of system

Types of systems[80]		
Single-parameter systems	Multiple-parameter systems	Aggregate weighted scores
ESI ED CIC	None identified	MEWS REMS mREMS PEDS RTS APACHE II THERM SCS MEES NEWS BEWS CCI MEWS plus NEWS-L SAPS II LODS MPMO II TEWS PIRO RAPS ASSIST VIEWS VIEWS-L PSI SOFA

The most common outcomes examined were in-hospital mortality (n=21), admission to Intensive Care Unit (ICU) (n=12), mortality (not specified where or during a specific follow up time frame possibly beyond hospital discharge) (n=10), hospital admission (n=7), and length of hospital stay (n=4).

Overall, the APACHE II score, PEDS, VIEWS-L, and THERM scores appeared relatively better at predicting mortality and ICU admission compared to other tools assessed in the included studies. The MEWS was the most commonly assessed tool and the cut-off value used was 4 or 5, with the exception of Dundar et al [79] who found an optimal cut-off of 4 for in-hospital mortality but 3 for predicting hospitalisation. For predicting ICU admissions, the AUROC of MEWS varied from 0.49 to 0.73 across studies. For the outcome in-hospital mortality, the AUROC ranged from 0.61 to 0.89. The BEWS contained the same parameters as the MEWS (respiratory rate, HR, SBP, temperature, LOC) and reported a 20% increase in risk of death within 48 hours and a 4% increase in risk for ICU admission with a BEWS score of five or more compared to a lower score.[69] The NEWS had a similar AUROC (0.70), predicting in-hospital mortality,[72] which is not surprising considering NEWS includes most of the same parameters as the MEWS. The MEWS plus score, which added the parameters age, race, gender, ED length of stay, method of arrival, and antibiotics given prior to or during ED visit, had a slightly greater AUROC (0.76) than MEWS.[54] The AUROC for the REMS ranged from 0.59 to 0.70 for ICU admission and 0.71 to 0.91 for in-hospital mortality.[59, 61, 67, 68] The modified REM score had an AUROC of 0.80[55] and VIEWS had an AUROC of 0.90[79] when predicting in-hospital mortality. The AUROC for predicting mortality was 0.64 for the CCI,[60] 0.71 for the PIRO score,[57] 0.71-0.90 for the APACHE II score,[57, 67] 0.65-0.87 for the RAPS,[67, 68] 0.69

for the MPMO II score, 0.72 for the SAPS II score, 0.60 for the LODS score,[56] and 0.83 for VIEWS-L.[73] The PEDS score had a higher AUROC for the prediction of death or admission to ICU (0.75-0.90) than the MEWS (0.73-0.76), the REMS (0.70), the APACHE II score (0.73), the RTS (0.75), the MEES (0.75), the NEWS (0.71) and the SCS (0.70).[61, 62] Cattermole et al [62] refined the PEDS score and developed the THERM score, which had an even higher AUROC (0.84) for prediction death or ICU admission.

Studies were subsequently categorised into three evidence tables according to the degree of differentiation of the ED patient group: a patient group with a certain (suspected) condition (Table 10), a patient group in a specific triage category(ies) (Table 11), or an undifferentiated patient group (Table 12).

Twelve of the 35 validation studies only included participants in (a) specific triage category(ies) (see Table 10 Evidence table below). Triage systems used varied across these studies, but included categories of patients that were critically ill and had to be seen with relative urgency (e.g. Manchester triage system I-III, Patient acuity category scale 1 or 2) or were admitted to the resuscitation room. Looking at the findings of this subgroup of studies in predicting mortality, the AUROC for the MEWS ranged from 0.63 to 0.75,[59, 64-66] it was 0.70-0.77 for REMS,[59, 61] 0.77-0.87 for NEWS,[76] 0.90 for PEDS, 0.83 for APACHE II, and 0.77 for RTS.[61] Predicting ICU admission, the AUROC were 0.54[59] and 0.49[64] for MEWS and 0.59 for REMS[59], while to predict hospital admission the AUROC of NEWS was 0.66-0.70.[76] Cattermole et al [61] and Cattermole et al [62] used a combined outcome of death and ICU admission and found an AUROC of 0.76 and 0.73 for MEWS, 0.90 and 0.75 for PEDS, 0.73 for APACHE II, 0.75 for RTS, 0.70 and 0.70 for REMS, 0.75 for MEES, 0.71 for NEWS, 0.70 for SCS, and 0.84 for THERM. One study assessed the prediction of septic shock by NEWS (AUROC 0.89).[50]

Eleven other studies (12 records) included a differentiated patient group with a specific (suspected) condition (see Table 11 Evidence table below). Five studies only included patients with (suspected) sepsis.[49, 52, 53, 57, 60, 71] Other study populations were restricted to patients with trauma,[73] suspected infection,[55, 77] pneumonia[72] or who had signs of shock.[56] Assessing the predictive ability of systems to predict mortality, MEWS had an AUROC of 0.61[60] and 0.72[52], CCI of 0.65,[60] mREMS of 0.80,[55] NEWS of 0.70,[72] NEWS-L of 0.73,[72] VIEWS-L of 0.83,[73] SAPS II of 0.72[56] and 0.90,[77] MPMO II of 0.69,[56] LODS of 0.60,[56] PIRO of 0.71,[57] APACHE II of 0.71[57] and 0.90,[77] and SOFA of 0.86.[77]

The remaining 12 studies assessed early warning systems in an undifferentiated ED population (see Table 12 Evidence table below). The AUROC to predict mortality was 0.71,[70] 0.73,[54] and 0.89[79] for MEWS, 0.76 for MEWS plus,[54] 0.91[67] and 0.85[68] for REMS, 0.87[67] and 0.65[68] for RAPS, and 0.90 for APACHE II.[67]
In summary, many different systems have been developed and evaluated to predict adverse outcomes in either differentiated or undifferentiated ED populations. The parameters most commonly included were HR, respiratory rate, blood pressure, temperature, oxygen saturations and level of consciousness. The MEWS was the most commonly assessed system and was better at predicting mortality than ICU admission, but the APACHE II score, PEDS, VIEWS-L, and THERM scores were relatively better at predicting mortality and ICU admission, although differences in study characteristics, the parameters measured and the weight given to individual parameters, may account for part of the observed differences in predictive ability.

Table 9. Evidence table: Development and validation studies – Scoping Review

(Methodological quality was rated using an adapted National Institute of Health checklist.[23] Full details of the methodological assessment are available in Appendix 3.)

Authors (year), country	Study aim	Study design	Setting & Participants	Content of system/ tool	Results
Challen & Goodacre (2011),[46] UK **Quality Rating**: Good	To carry out a scoping review of the literature relating to outcome prediction in adult non-trauma emergency patients, in order to identify the number and range of risk scores developed for acutely ill adults and to identify the outcomes these scores predict.	Scoping review	Papers that detailed a clinical assessment tool that was applied at the point of patient presentation to unscheduled healthcare services with outcome measures 30 days after presentation.	Selection criteria: Wholly or predominately clinical assessment tool, adult population, an outcome measure up to 30 days after presentation.	<u>Tools:</u> Scoring systems available for 17 broad conditions (with 80 different inclusion criteria) 119 tools assessed (of which 25 generic) <u>Outcomes:</u> 51 different outcome measures used (of which 30 disease specific) Analyses that used 'death' as outcome (247): 190 reported AUROC of which 69 AUROC > 0.8. Analyses that did not use 'death' as outcome (251): 151 reported AUROC of which 30 AUROC >0.8 Lowest AUROC = 0.44 (predicting hospital deaths in patients with acute MI) Highest AUROC = 0.98 (APACHE II for predicting hospital deaths in patients with peritonitis)

Table 10. Evidence table: Development and validation studies – Patient groups differentiated by triage category

(Risk of bias was rated using a tool adapted from Kansagara et al (2011).[30] Full details of the risk of bias assessment are available in Appendix 3.)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Alam et al (2015),[76] the Netherlands **Risk of bias:** Unclear	To explore the performance of NEWS with regard to predicting adverse outcomes in adult patients and the ability of NEWS to predict the need for hospital admission in an ED population.	Prospective cohort (Validation)	274 patients (≥18 years) presenting (T0) to the ED of an urban academic tertiary care centre between 7 Jan-15 Feb 2013 with an Emergency Severity Index score of 2 and 3 not triaged to the resuscitation room. For 247 of these 274 patients, the NEWS was calculated an hour later (T1). Only 133 of the 247 patients could be followed up at discharge from the ED (T2).	**NEWS**: Parameters: Respiratory rate, SBP, HR, temperature, oxygen saturation	Hospital admission, length of stay, ICU admission, mortality	**Hospital admission** (n=130) NEWS significantly associated with admission at all 3 time points ($p < 0.001$). • T0: AUROC 0.66 (95% CI 0.60–0.73) • T1: AUROC 0.69 (95% CI 0.62–0.75) • T2: AUROC 0.70 (95% CI 0.61–0.79) **Length of stay** NEWS significantly associated with length of stay at all 3 time points ($p < 0.001$). Median length of stay more than doubled for a score >7 compared with a score of 0–4. (AUROC not provided) **ICU admission** (n=10) NEWS significantly associated with ICU admission at all 3 time points (T0: p=0.003; T1: p=0.001; T2: p=0.046). (AUROC not provided) **30-day Mortality** (n=11) NEWS significantly associated with mortality at all 3 time points ($p < 0.001$). 30-day mortality was not significantly related to ESI ($p = 0.816$). • T0: AUROC 0.77 (95% CI 0.62–0.92) • T1: AUROC 0.87 (95% CI 0.77–0.96) • T2: AUROC 0.77 (95% CI 0.57–0.97). Of the individual physiological measures of NEWS: • Respiratory rate was associated with mortality at all measured time points (T0: p=0.017; T1: p<0.001; T2: p=0.014). • Pulse rate had a strong correlation with mortality at T1 (p=0.037) No correlations could be found for all other physiological parameters.

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Armagan et al (2008),[58] Turkey **Risk of bias**: Unclear	To determine the predictive validity of the Modified Early Warning Score (MEWS) in a Turkish ED setting.	Prospective cohort (Validation)	309 patients (Triage I, II, II) in ED of one hospital between April-Aug 2007.	**MEWS:** Parameters: SBP, pulse rate, respiratory rate, temperature, AVPU score	Hospital admission, ICU admission, in-hospital death, ED death	Low risk (MEWS≤4) (n=106); high risk (MEWS >4) (n=203) MEWS (cut-off >4) • Admission to hospital: adjusted OR 1.56 (95% CI 0.93-2.98) • Admission to ICU: adjusted OR 1.95 (95% CI 1.04-366.00) (p=0.04) • Death in ED: adjusted OR 35.13 (95% CI 4.58-269.40) (p<0.001) • Death in hospital: adjusted OR 14.80 (95% CI 5.52-39.70) (p<0.001)
Bulut et al (2014),[59] Turkey **Risk of bias**: Low	Compare the efficacy of Modified Early Warning Score (MEWS) and Rapid Emergency Medicine Score (REMS) on in-hospital mortality, and as a predictor of hospitalisation in general medical and surgical patients admitted to ED.	Prospective, multicentre cohort (Validation)	2000 general medical & surgical patients (red & yellow triage category) presenting to EDs of 3 hospitals between Oct 2011-April 2012.	**REMS:** Parameters: Age, HR, Temp, Respiratory Rate, Mean arterial pressure, GCS, oxygen saturations **MEWS:** Parameters: SBP, HR, Respiratory rate, Temp, AVPU	Admission to ward or ICU/HDU, in-hospital mortality	Median (range): • MEWS: 1 (0-9); REMS: 5 (0-16) 40.8% hospitalised ward, 29.8%-ICU/HDU, 29.2% discharged. Total in-hospital mortality was 7.7% (n=153). **Predicting in-hospital mortality** • REMS (6–13) vs REMS <6: RR 2.92 (95% CI 0.03 to 4.22); p<0.001 • REMS (>13) vs REMS <6: RR 14.56 (95% CI 4.57 to 46.57); p<0.001 • MEWS≥5 vs MEWS <5: RR 3.84 (95% CI 2.36 to 6.24); p<0.001 • MEWS AUROC: 0.63 (95% CI 0.61-0.65) • REMS AUROC: 0.71 (95% CI 0.67-0.72) • Performance of REMS was higher (p<0.001) **Predicting discharge vs hospitalisation:** • MEWS AUROC: 0.57 (95% CI 0.55-0.59) • REMS AUROC: 0.64 (95% CI 0.62-0.66) • Performance of REMS was higher (p<0.001) **Predicting admission to ICU/HDU:** • MEWS AUROC: 0.54 (95% CI 0.52-0.56) • REMS AUROC: 0.59 (95% CI 0.57 to 0.61) • Performance of REMS was higher (p<0.001)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Cattermole et al (2009),[61] Hong Kong **Risk of bias**: Low	(1) Determine a new prognostic score making use of rapidly available and easily measurable physiological parameters and initial laboratory tests in resuscitation room patients, in order to identify patients most at risk of death or in need of ICU care. (2) To compare the new score with APACHE II, RTS, REMS and MEWS scores	Prospective cohort (Development & Validation)	330 ED patients ≥18 years of age triaged to resuscitation room in 1 hospital between 9th April & 6th May 2006.	**Prince of Wales ED Score (PEDS)** (new score) Parameters: SBP, GCS, Glucose, HCO_3, white blood cells, metastatic cancer history **Revised Trauma Score** (RTS) **Rapid Emergency Medicine Score** (REMS) MEWS **Acute Physiology and Chronic Health Evaluation score** (APACHE II) Parameters not stated in this report	**Primary outcome**: death admission to ICU within 7 days of ED attendance (vs survival at 7 days without ICU admission) **Secondary outcome measures:** 30 day mortality and hospital length of stay.	Poor outcome: 23.0% (77/330) (40 (12.1%) admitted to ICU or 41 (12.4%) died within 7 days). PEDS score ranged from -2 to 58. **Comparison of PEDS, APACHE II, RTS, REMS and MEWS (n=234) for the primary outcome:** PEDS: AUROC 0.90 (0.87–0.94), Sensitivity 0.87, Specificity 0.80, PPV 0.57, NPV 0.95, Accuracy 0.82 APACHE II: AUROC 0.73 (0.68–0.78), Sensitivity 0.61, Specificity 0.70, PPV 0.44, NPV 0.86 Accuracy 0.73 RTS: AUROC 0.75 (0.70–0.79), Sensitivity 0.58, Specificity 0.83, PPV 0.51, NPV 0.87 Accuracy 0.82 REMS: AUROC 0.70 (0.64–0.75), Sensitivity 0.51, Specificity 0.79, PPV 0.42, NPV 0.84 Accuracy 0.73 MEWS: AUROC 0.76 (0.71–0.81), Sensitivity 0.69, Specificity 0.74, PPV 0.45, NPV 0.89 Accuracy 0.73 **Comparison of PEDS, APACHE II, RTS, REMS and MEWS (n=234) for the secondary outcome (30 day mortality):** PEDS: AUROC 0.90 (0.86–0.93) APACHE II: AUROC 0.84 (0.79–0.88) RTS: AUROC 0.77 (0.72–0.81) REMS: AUROC 0.77 (0.72–0.82) MEWS: AUROC 0.75 (0.70–0.80)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Cattermole et al (2013),[62] Hong Kong **Risk of bias**: Unclear	(1) To validate Prince of Wales ED Score (PEDS) in comparison with other prognostic scores (MEWS, SCS, REMS, MEES, MEDS, Worthing & NEWS); (2) to simplify and refine the score, using only variables that are immediately available in the resuscitation room; (3) to validate the new score to devise original PEDS	Prospective cohort (Development & Validation)	234 consecutive, ≥18 years of age, patients managed in resuscitation room during weekdays over a 3 month period.	**The Resuscitation Management score** (THERM) (refined PEDS score): Parameters: GCS, HCO_3^- and SBP **PEDS** Parameters: SBP, GCS, Glucose, HCO_3, white blood cells, metastatic cancer history **MEWS** **Simple Clinical Score** (SCS) **REMS** **Mainz Emergency Evaluation Score** (MEES) **National Early Warning Score** (NEWS) **Worthing** (excluded because it is a condition specific score) **Mortality in the ED Sepsis** (MEDS) (excluded because it is a condition specific score) Parameters not stated in report.	**Primary outcome**: death or admission to ICU. **Secondary outcomes**: 30 day mortality & hospital length of stay.	37/234 admitted to ICU or died within 7 days. PEDS: AUROC 0.75 (95% CI 0.69 to 0.80) MEES: AUROC 0.75 (95% CI 0.69 to 0.80) MEWS: AUROC 0.73 (95% CI 0.67 to 0.79) NEWS: AUROC 0.71 (95% CI 0.64 to 0.76) REMS: AUROC 0.70 (95% CI 0.64 to 0.76) SCS: AUROC 0.70 (95% CI 0.64 to 0.76) THERM scores: max score=37, High risk <30, medium risk (30.1-35), low risk (35.1-37). **Comparison of THERM and NEWS (n=234)** THERM: AUROC: 0.84 (0.79 to 0.88) High risk cut-off: Sensitivity 0.57 (0.40 to 0.73), Specificity 0.89 (0.84 to 0.93), PPV 0.50 (0.34 to 0.66), NPV 0.92 (0.87 to 0.95) Medium risk cut-off: Sensitivity 0.89 (0.75 to 0.97), Specificity 0.65 (0.58 to 0.72), PPV 0.32 (0.23 to 0.42), NPV 0.97 (0.92 to 0.99) NEWS: AUROC: 0.71 (0.64 to 0.76) High risk cut-off: Sensitivity 0.65 (0.48 to 0.80), Specificity 0.71 (0.64 to 0.77), PPV 0.29 (0.20 to 0.40), NPV 0.91 (0.86 to 0.95) Medium risk cut-off: Sensitivity 0.92 (0.78 to 0.98), Specificity 0.44 (0.37 to 0.51), PPV 0.24 (0.17 to 0.31), NPV 0.97 (0.91 to 0.99) THERM had superior specificity; there was no significant difference in AUROC, sensitivity or predictive values.

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Christensen et al (2011),[69] Denmark **Risk of bias**: Low	To evaluate ability of the Bispebjerg EWS (BEWS) to identify critically ill patients in ED and to examine feasibility of using BEWS to activate Medical Team response	Retrospective cohort (Development & Validation)	Random sample of 300 'red' category patients visiting ED of 1 hospital between April-Sept 2009.	**BEWS** Parameters: Respiratory Rate, HR, SBP, Temp, LOC	Admission to ICU within 48 hrs of arrival at ED or death within 48hrs of arrival at ED.	138 patients out of 300 were excluded for insufficient data; 162 included. Activated Emergency Call (EC): 24 Admitted to ICU within 48 hrs of ED: 4 (2 died) Deaths within 48hrs ED: 6 BEWS ≥ 5 (vs <5): **Death within 48 hours of arrival:** • RR 20.3 (95% CI 6.9-60.1) • Sensitivity 83.0%, Specificity 83.0%, PPV 16.0, NPV 99.0 ICU admission within 48 hours of arrival: • RR 4.1 (95% CI 1.5- 10.9) • Sensitivity 50.0%, Specificity 81.0%, PPV 6.0, NPV 98.0 **Critically ill:** • RR 6.8 (95% CI 3.3-13.8) • Sensitivity 63.0%, Specificity 82.0%, PPV 16.0, NPV 98.0

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Gu et al (2015),[32] China **Risk of bias:** Unclear (Only abstract in English)	To evaluate the value of Modified Early Warning Score (MEWS) in predicting mortality of critically ill patients admitted to emergency department.	Prospective cohort (Validation)	176 emergency patients admitted to resuscitation room of one hospital between 13 Feb-20 April 2014.	**MEWS** (parameters not listed in abstract)	3-day mortality, all deaths, and composite outcome of intensive care unit (ICU) transfer, cardio-pulmonary resuscitation, and death.	Mean MEWS 4.30±2.74; 74 cases MEWS ≥ 5 and 102 in MEWS 0-4. **3-days mortality (n=41)** MEWS 0-1 (12.7% (13/102); ref) MEWS ≥ 5 (37.8 (28/74); OR 4.2 (95%CI 2.0 - 8.8, P < 0.001)) Multi-regression logistic showed abnormal mental status (OR 3.6, 95% CI = 1.5-8.4, P = 0.003) but not MEWS ≥ 5 (OR = 1.7, 95%CI = 0.6-4.5, P = 0.3) was the predictor of 3-day mortality. **All death (n=58)** MEWS 0-1 (17.7% (18/102); ref) MEWS ≥ 5 (54.1 (40/74); OR 5.5 (95%CI 2.8- 10.9, P < 0.001)) **ICU transfer, cardio-pulmonary resuscitation and death (n=74)** MEWS 0-1 (25.5% (26/102); ref) MEWS ≥ 5 (64.9 (48/74); OR 5.4 (95%CI 2.8- 10.4, P < 0.001))

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Ho et al (2013),[64] Singapore **Risk of bias**: Low	To validate the use of the Modified Early Warning Score (MEWS) as a predictor of patient mortality and intensive care unit (ICU)/ high dependency unit (HDU) admission in an Asian population.	Retrospective cohort (Validation)	1024 critically ill patients, ≥18 years of age, presenting to a large Asian tertiary ED between Nov 2006 and Dec 2007, and requiring continuous ECG monitoring and Patient Acuity category Scale (PACS) of 1 or 2.	**MEWS:** Parameters: SBP, pulse rate, respiratory rate, temperature, AVPU score	Mortality during the inpatient period following admission from the ED up to 30 days, and direct admission from the ED to the high dependency unit, intermediary care area or the intensive care unit.	713 patients MEWS score <4; 311 patients MEWS ≥4. **Mortality** 47 deaths (6.6%) in MEWS <4; 53 (17.0%) deaths in MEWS ≥4 (p<0.001) For cut-off value ≥4: Sensitivity: 47.0 Specificity: 27.9 PPV: 6.7 NPV: 83.0 AUROC: 0.68 **Admission** 267 (37.4%) were admitted to HDU/ICA in MEWS <4; 86 (27.7%) admitted to HDU/ICA in MEWS ≥4 (p=0.00). For cut-off value of 4: Sensitivity: 74.2 Specificity: 33.9 PPV: 46.7 NPV: 62.7 AUROC: 0.5

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Hock Ong et al (2012),[65] Singapore **Risk of bias**: Unclear	To validate a novel Machine Learning (ML) score incorporating Heart Rate Variability (HRV) for risk stratification of critically ill patients presenting to the ED by comparing the area under the curve, sensitivity and specificity for prediction of cardiac arrest with the Modified Early Warning Score (MEWS).	Prospective cohort (Validation)	925 patients, ≥18 years of age, requiring continuous ECG monitoring triaged as Patient Acuity Category Scale (PACS) 1 or PACS 2 were eligible.	**MEWS:** Parameters: SBP, pulse rate, respiratory rate, temperature, AVPU score **ML-based score**: (excluded because HRV is not a routine measure)	Cardiac arrest within 72 hours of presentation to the ED, death after admission (in-hospital death during current admission, including within 72 hours).	4.6% (43) developed cardiac arrest within 72 hours; 9.3% (86) died after admission. **Cardiac arrest** **MEWS** Sensitivity: 74.4 Specificity: 54.2 PPV: 7.4 (5.3-10.3) NPV: 97.8 (95.9-98.8) +LR: 1.6 (1.3-2.0) AUROC: 0.7 **Death after admission** **MEWS** Sensitivity: 74.4 Specificity: 55.7 PPV: 14.7 (11.5-18.4) NPV: 95.5 (93.2-97.1) +LR: 1.7 (1.5-1.9) AUROC: 0.7
Keep et al (2015),[50] UK **Risk of bias**: Low	Explore relationship between initial National Early Warning Score (NEWS UK) in ED and diagnosis of Severe Sepsis (SS)	Retrospective cohort (Validation)	500 patients, >16 years of age, presenting at ED of an urban hospital in a 5-day period, with triage category 1-3 (Manchester Triage Score).	**NEWS:** Parameters: Respiratory Rate, HR, O_2 saturations, SBP, Temp, LOC.	Septic shock (SS)	Sepsis: 9.9% (n=50); Septic shock (SS): 5.4% (n=27) **Prediction of SS:** NEWS optimal cut-off ≥3: AUC 0.90 (95% CI 0.84-0.94). Sensitivity 92.6% (95% CI, 74.2%-98.7%), specificity of 77% (95% CI 72.8% to 80.6%), PPV 18.7% (95% CI 12.7% to 26.5%), NPV 99.5% (95%CI 97.8%-99.9%)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Lui et al (2014),[66] Singapore **Risk of bias**: Unclear	To propose an intelligent scoring system and explore the utility of combining Heart Rate Variability (HRV) and 12-lead ECG parameters, and vital signs to predict acute cardiac complications within 72 h.	Prospective cohort (Development & Validation)	564 (of eligible 702) chest pain patients aged ≥30, triaged as PACS 1 or 2, recruited between March 2010-April 2012 at the ED of 1 hospital.	**ESS** (Proposed Ensemble-Based Scoring System: combines HRV and 12-lead ECG parameters, and vital signs). (Excluded because not routine measurement) **DIST** (an Euclidean distance-based scoring system): uses HRV combined with vital signs (Excluded because not routine measurement) **MEWS** (modified early warning score: reference to Subbe et al 2003) **TIMI** (thrombolysis in myocardial infarction: reference to Antman et al 2000) (Excluded because condition (MI) specific score)	Composite of four severe complications within 72 h of arrival at the ED: mortality, cardiac arrest, sustained ventricular tachycardia, and hypotension requiring inotropes or intraaortic balloon pump insertion.	19 (3.4%) out of the remaining 564 patients met the composite outcome. **MEWS** Cut-off: 1.0 AUROC: 0.67 (0.54-0.81) Sensitivity: 42.1% (19.9%-64.3%) Specificity: 78.5% (75.1%-82.0%) PPV: 6.4% (2.1%-10.7%) NPV: 97.5% (96.0%-99.9%)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Wilson et al (2016),[43] UK **Risk of bias**: High	To evaluate the ability of a data-fusion Patient Status Index (PSI) to detect patient deterioration in the ED in comparison with documented TTS and retrospectively calculated TTS.	Prospective cohort (Validation)	472 adults (≥18) entering one of three clinical areas of the ED (resuscitation room, 'majors', observation ward); during times the research team was available (daytime).	**Vital sign and TTS data**: heart rate (HR), systolic and diastolic blood pressure (BP), respiratory rate, peripheral oxygen saturation (SpO_2), temperature and Glasgow Coma Scale (GCS) score. **Paper TTS**: clinician assigns score **eTTS**: For each set of manually recorded vital signs, TTS was retrospectively calculated **PSI**: reference to Tarassenko et al (2006); parameters: HR, respiratory rate, BP, temperature, oxygen saturation	Escalation of care	**PSI true alerts** Escalation after ED arrival: 35, of which 20 had PSI data, of which: **1.** Detected by TTS: 4 **2.** Detected by eTTS: 17 **3.** Detected by PSI: 15 **4.** Detected by eTTS, not PSI: 5 **5.** Detected by PSI, not eTTS: 3 **PSI false alerts** False alert rate: 1.13 alerts/bed-day (49 false alerts from 39 patients).

Table 11. Evidence table: Development and validation studies – Patient groups differentiated by (suspected) condition

(Risk of bias was rated using a tool adapted from Kansagara et al (2011).[30] Full details of the risk of bias assessment are available in Appendix 3.)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Albright et al. (2014),[53] USA **Risk of bias**: Low	To design an emergency department sepsis scoring system to identify risk of ICU admission in pregnant & postpartum women.	Retrospective cohort (Development & validation)	850 pregnant & post-partum women with suspected SIRS or sepsis evaluated in ED of a large tertiary care hospital between Feb 2009-May 2011.	**Sepsis in Obstetrics Score (SOS)**: (Excluded because it is a condition (sepsis) specific tool) **MEWS** (Parameters not stated in report.) **REMS** (Parameters not stated in report.)	**Primary outcome**: ICU admission within 48 hrs of ED presentation. **Secondary outcomes:** • telemetry unit admission • length of hospital stay • mortality • positive blood cultures • positive influenza swab • antibiotic use • adverse perinatal outcome	9 (1.1%) admitted to ICU, 32 (3.8%) to telemetry unit, mortality (0.0%). **Primary outcome: ICU Admission prediction** MEWS (cut-off ≥5): Sensitivity: 100.0%, Specificity: 77.6%, PPV: 4.6%, NPV: 100.0% REMS (cut-off ≥6): Sensitivity: 77.8%, Specificity: 93.3%, PPV: 11.1%, NPV: 99.7% **Secondary outcomes** Not reported for MEWS and REMS

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Cildir et al (2013),[60] Turkey **Risk of bias:** Low	To evaluate the modified Mortality in Emergency Department Sepsis (MEDS) score, MEWS score and CCI to predict prognosis in patients presenting to ED diagnosed with sepsis.	Prospective cohort (Validation)	230 patients ≥18 who presented to ED of 1 hospital between 7 Aug 2009-15 Feb 2011, diagnosed with community acquired sepsis.	**Charlson Comorbidity Index (CCI)**: parameters not stated in report **MEWS:** parameters not stated in report **Mortality in emergency department sepsis (MEDS)**: excluded from review because it is condition (sepsis) specific	Mortality	**Prediction of the mortality in the group with sepsis according to 28-day mortality (n = 64) (8 patients died)** CCI (cut-off >5) Sensitivity 50%, Specificity 85.4%, PPV 33.3, NPV 92.2, AUC 0.65 (p=0.18) MEWS (cut-off ≤5) Sensitivity 87.5%, Specificity 30.4%, PPV 15.2, NPV 94.4, AUC 0.57 (p=0.48) **Prediction of the mortality in the group with severe sepsis according to 28-day mortality (n=166) (66 patients died)** CCI (cut-off >5) Sensitivity 78.8%, Specificity 38%, PPV 45.6, NPV 73.1, AUC 0.62 (p=0.006) MEWS (cut-off ≤5) Sensitivity 48.5%, Specificity 67.0%, PPV 49.2, NPV 66.3, AUROC 0.60 (p=0.04) **Predictive value of the scores for 28-day mortality** CCI (cut-off >5) Sensitivity 27.0%, Specificity 93.0%, PPV 64.5, NPV 72.9, AUROC 0.65 (p=0.001) MEWS (cut-off ≤5) Sensitivity 43.2%, Specificity 75.0%, PPV 45.1, NPV 73.6, AUROC 0.61 (p=0.008)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Considine et al (2015),[78] Australia	To evaluate the effect of the staged implementation of a Rapid Response System on reporting of clinical deterioration in ED patients. A secondary aim was to determine if there were differences between patients who did, and did not, experience clinical deterioration documented during ED care.	Retrospective cross-sectional design (Validation)	Stratified random sample of 600 adult ED patients (≥18 years) with presenting complaints of shortness of breath, chest pain or abdominal pain in a 300-bed urban district hospital. 150 patient in each of 4 groups. Four groups according to stage of implementation: (T0) Clinical decision making/ discretion, no TTS chart (year 2009); (T1) Escalation of care protocol (if any critical instability criteria met, an emergency physician should review within 5min), no TTS chart (year 2010); (T2) Escalation of care protocol, single parameter TTS chart (year 2011); (T3) Escalation of care protocol, single parameter TTS chart (year 2012).	**Critical instability criteria (ED CIC):** • Airway/ breathing: Stridor, upper airway obstruction, or threatened airway, SpO2 < 90% (on oxygen 10 L/min via mask), Arterial blood gases pH < 7.20, Respiratory rate < 10 breaths/min or > 30 breaths/min • Circulation: Heart rate < 50 beats/min or > 120 beats/min, Systolic blood pressure < 90 mmHg or > 200 mmHg, Urine output < 20 mL/h or < 100 mL/6 h • Disability: Sudden decrease in consciousness (fall in Glasgow Coma Scale score > 2), Repeated or prolonged seizures • Worried?: Patients who may not meet the above criteria but have a sudden deterioration in their medical condition, requiring urgent medical review.	**Primary outcome**: Unreported clinical deterioration (=presence of documented physiological abnormalities that fulfilled the ED Clinical Instability Criteria in ED nursing notes and no documentation that these were reported to a medical officer.)	At T0, 86.7% of episodes of clinical deterioration were unreported. Across the four years studied, episodes of unreported clinical deterioration decreased by 17.9% from T0 to T1 (68.8%), 13.5% from T1 to T2 (55.3%) but only 1.3% from T2 to T3 (54.0%); none of these differences were statistically significant ($p = 0.14$). Patients who experienced clinical deterioration in the ED were more likely to arrive by ambulance ($p < 0.001$), be triaged to Australian Triage Scale categories 1 or 2, ($p < 0.001$), and had a 2.8 hour longer median ED length of stay, and were 31.9% more likely to be admitted to hospital ($p < 0.001$).

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Corfield et al (2014)[81] (and related conference abstract Corfield et al (2012),[49] Scotland **Risk of bias**: Low	To determine whether a single National Early Warning Score (NEWS) on ED arrival is a predictor of outcome in patients with sepsis, either in-hospital death in 30 days or ICU admission within 2 days	Retrospective cohort (Validation)	3890 (74% of 5285 eligible patients) adult patients, >16 years of age, attending ED with sepsis (suspected or confirmed within 2 days of attendance and 2 or more of sepsis criteria). 20/25 Scottish mainland EDs participated.	NEWS (0-20 score) Parameters: respiratory rate, oxygen saturations, temperature, SBP, pulse, conscious level, supplemental O_2	ICU admission within 2 days of attendance at ED and 30-day mortality (in-hospital). A combined endpoint of ICU admission/ and or mortality was also assessed.	Included in analysis: n = 2003 ICU (within 2 days): n = 113 (6.0%) 30-day mortality: n = 297 (15.0%) Combined (ICU and/or mortality): n = 376 (19.0%) **ICU (within 2 days)** (Compared to NEWS 0-4; Adjusted for age) NEWS Score 5-6: OR 1.22 (95% CI 0.59-2.54; p=0.59) 7-8: OR 2.01 (95% CI 1.02-3.97; p=0.04) 9-20: OR 5.76 (95% CI 3.22-10.31; p=0.00) **Mortality (30 days)** (Compared to NEWS 0-4; Adjusted for age) NEWS Score 5-6: OR 1.95 (95% CI 1.21-3.14; p=0.01) 7-8: OR 2.26 (95% CI 1.42-3.61; p=0.004) 9-20: OR 5.64 (95% CI 3.70-8.60; p=0.00) **Combined (ICU and/or mortality)** (Compared to NEWS 0-4; Adjusted for age) NEWS Score 5-6: OR 1.72 (95% CI 1.14-2.60; p=0.01) 7-8: OR 2.17 (95% CI 1.45-3.25; p=0.00) 9-20: OR 5.78 (95% CI 4.02-8.31; p=0.00) Cut-off point with highest Youden's Index[1]: NEWS 9: Sensitivity 0.52, specificity 0.77, PPV 0.35, NPV 0.88, Youden's index 0.30

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Geier et al (2013),[71] Germany **Risk of bias**: Low	(1) To evaluate Emergency Severity Index (ESI), Modified Early warning Score (MEWS), Mortality in Emergency Department Sepsis (MEDS) score concerning their diagnostic accuracy to detect patients with Severe Sepsis and Septic Shock score (SSSS). (2) To determine the prognostic accuracy of indices in predicting the in-hospital mortality of patients with suspected sepsis in ED. (3) To calculate the prognostic value of the Charlson Comorbidity Index (CCI).	Prospective cohort (Development & Validation)	151 consecutive adult patients with suspected sepsis admitted to ED of 1 hospital between 1 Aug-30 Sept 2012.	**Emergency Severity Index (ESI)** (5 levels; the higher the level the lower the medical urgency) Level 1 = acute life threatened ill patients - require immediate initiation of diagnostics and therapy. Level 2 = patients in high-risk situation - initiation of diagnostics and therapy has to start within 10min following the initial triage assessment. Higher levels not specified in report. **MEWS** (0-14) Parameters: SBP, HR, Temperature, respiratory rate, LOC **CCI Score** (0-37) Parameters not specified in report. **MEDS score** (Excluded because it is a condition (sepsis) specific score).	In-hospital mortality	45.0% (n=72) diagnosed with SSSS; 33.1% (n=53) uncomplicated sepsis (without organ dysfunction). 21.9% (n=26) no sepsis, but SIRS or locally confined infection. **In-hospital mortality** (14.6% of all patients; 27.8% of patients with SSSS) ESI Sensitivity 0.73, Specificity 0.0, PPV 0.17, NPV 0.90 MEWS Sensitivity 0.43, Specificity 0.74, PPV 0.21, NPV 0.89 CCI Sensitivity 0.82, Specificity 0.64, PPV 0.21, NPV 0.94

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Howell et al (2007),[55] USA **Risk of bias**: Low	To validate the Mortality in Emergency Department Sepsis (MEDS) score, the Confusion, Urea nitrogen, Respiratory rate, Blood pressure, 65 years of age and older (CURB-65) score, and a modified Rapid Emergency Medicine Score (mREMS) in patients with suspected infection.	Prospective cohort (Validation)	2132 adult patients with clinically suspected infection admitted to an urban ED between 10 Dec 2003 - 30 Sept 2004.	**mREMS** points depend on severity Mean arterial pressure (mm Hg) Pulse rate Respiratory rate Peripheral oxygen saturation Glasgow Coma Score Age **MEDS score** (Excluded because it is a condition (sepsis) specific score.) **CURB-65 score** (Excluded because it is a condition (community-acquired pneumonia) specific score.)	28-day in-hospital survival (patients discharged alive from the hospital before 28 days were considered alive for the 28-day in-hospital mortality end point).	Of 2,132 patients with unique first visits, 83 (3.9%; 95% CI 3.1% to 4.7%) died. **mREMS** Odds of death increased by 1.40 (95% CI = 1.28 to 1.45) with each point increase. AUROC 0.80 (0.75-0.85)
Jo et al (2013),[73] Korea **Risk of bias**: Low	To compare the predictive value of the VitalPAC Early Warning Score-Lactate (VIEWS-L) score with that of the Trauma Injury Severity Score (TRISS).	Retrospective cohort (Validation)	299 patients, ≥ 15 years of age, with blunt trauma, Injury severity score ≥9 in a 1000-bed urban hospital between 1 Apr 2010-31 March 2011.	**VIEWS-L** Parameters: SBP, HR, respiratory rate, temperature, oxygen saturations, inspired oxygen, central nervous system alertness **TRISS** (Excluded because it is a trauma-specific tool)	In-hospital mortality	**VIEWS-L** Higher score in non-survivors (median 7.9, IQR 6.2-12.9) than non-survivors (median 3.7, IQR 1.7-5.5). AUROC: 0.83 (95% CI 0.77-0.91)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Jo et al. (2016),[72] Korea **Risk of bias**: Low	To investigate the prognostic prediction power of newly introduced early warning score modified by serum lactate level, the National Early Warning Score (UK) including Lactate (NEWS-L), among community-acquired pneumonia patients, and compared with previously used tools such as Pneumonia Severity Index and CURB-65.	Retrospective cohort (Validation)	553 patients, ≥18 years of age, with an admission diagnosis of any type of pneumonia between 1 Oct 2013-30 Sept 2014 in 1 hospital.	**NEWS-L score** Parameters: SBP, HR, respiratory rate, temperature, SpO_2, LOC, supplemental oxygen, lactate level **CURB-65** (Excluded because it is a condition (community-acquired pneumonia) specific tool) **Pneumonia Severity Index** (Excluded because it is a condition (pneumonia) specific tool)	Inpatient mortality	Mortality by NEWS-L score: ≤ 3.0: 2.2% 3.1 ≤ and ≤ 5.2: 7.9% 5.3 ≤ and ≤ 8.0: 9.6% ≥8.1: 23.9% **NEWS-L** AUROC; 0.73 (0.66-0.80); reference Cut-off ≥3.1: Sensitivity 95.0 (86.1-99.0) Specificity 27.6 (23.7-31.8) PPV (%) 13.8 (10.6-17.5) NPV (%) 97.8 (93.8-99.6) PLR 1.3 (1.2-1.4) NLR: 0.2 (0.06-0.6) Cut-off ≥5.3: Sensitivity 76.7 (64.0-86.6) Specificity 53.8 (49.2-58.2) PPV 16.8 (12.6-21.8) NPV 95.0 (91.7-97.2) PLR 1.7 (1.4-2.0) NLR 0.4 (0.3-0.7) Cut-off ≥8.1: Sensitivity 55.0 (41.6-64.9), Specificity 78.7 (74.8-82.2), PPV 23.9 (17.1-31.9), NPV 93.5 (90.7-95.7), PLR 2.6 (1.9-3.4), NLR 0.6 (0.4-0.8) Cut-off ≥7.3: Sensitivity 63.3 (49.9-75.4), Specificity 73.2 (69.1-77.1), PPV 22.4 (16.3-29.4), NPV 94.3 (91.4-96.4), PLR 2.4 (1.9-3.0), NLR 0.5 (0.4-0.7) **NEWS** AUROC 0.70 (0.63-0.77); p=0.03 Cut-off ≥5 including red score: Sensitivity 68.3 (54.9-79.4), Specificity 57.2 (52.7-61.6), PPV 16.3 (12.1-21.5), NPV 93.7 (90.2-96.1), PLR 1.60 (1.3-2.0), NLR 0.55 (0.4-0.8)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Jones et al (2005),[56] USA **Risk of bias**: Low	The hypothesis of this study states that the physiologic scoring systems New Simplified Acute Physiology Score II (SAPS II), Morbidity Probability Model (MPM0 II), and Logistic Organ Dysfunction System (LODS), when calculated from variables available during the care of critically ill ED patients, will not perform with high diagnostic accuracy for predicting hospital mortality, defined as an area under the receiver operating characteristic (ROC) curve <0.80.	Secondary analysis of a randomized controlled trial (Validation)	91 (45% of eligible patients) non-trauma ED patients admitted to an intensive care unit, >17 years of age, with initial ED vital signs consistent with shock (systolic blood pressure <100 mm Hg or shock index >1.0), and with agreement of two independent observers for at least one sign and symptom of inadequate tissue perfusion.	**SAPS II** **MPM0 II** **LODS** No data provided on included parameters.	In-hospital mortality	The in-hospital mortality rate was 21% (19/91). **SAPS II** Mean 40 (SD 14) Predicted mortality 28% AUROC 0.72 (0.57-0.87) **MPM0 II** Mean -1.06 (SD 1.24) Predicted mortality 28% AUROC 0.69 (0.54-0.84) **LODS** Mean 5 (SD 3) Predicted mortality 30% AUROC 0.60 (0.45-0.76) Using only ED variables to calculate the scores resulted in all three scoring systems overestimating in-hospital mortality by a mean of 8% (range, 7–10%). The scoring systems appear to function most accurately in the lower risk group for each scoring system (SAPS II ≤50, LOD ≤7, and MPM0 II ≤0) with an average difference between actual and predicted mortality of only 3%. All three of the scoring systems greatly overestimated mortality in the higher risk group, with an average difference between actual and predicted mortality of 31%.

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Nguyen et al (2012),[57] USA **Risk of bias:** Unclear	To examine the performance of the Predisposition, Insult/Infection, Response, and Organ dysfunction (PIRO) model compared with the Acute Physiology and Chronic Health Evaluation (APACHE) II and Mortality in Emergency Department Sepsis (MEDS) scoring systems in predicting in-hospital mortality for patients presenting to the emergency department (ED) with severe sepsis or septic shock.	Prospective cohort (Validation)	541 patients, >17 years of age, who met the criteria for high-R severe sepsis (sepsis with lactate ≥4 mmol/L) or septic shock in the ED.	**PIRO** Parameters: Age, chronic liver disease, and/or congestive cardiomyopathy Infection Tachycardia/tachyopnea Organ dysfunction **APACHE II** reference Knaus et al (1985) **MEDS** (Mortality in Emergency Department Sepsis; reference Sapiro et al 2003) (Excluded because it is a condition (sepsis) specific tool)	In-hospital mortality	62% (61.9) of patients were diagnosed with septic shock; 63.4%, with positive culture; and 46.9%, with positive blood culture. During the course of care, 31.8% patients died in the hospital. **PIRO** Predicted mortality: 48.5% (40.1 and 63.9) AUROC: 0.71 (0.66-0.75) **APACHE II** Predicted mortality: 66% (42 and 83) AUROC: 0.71 (0.66-0.76) Actual mortality significantly increased with increasing PIRO score in patients with either APACHE II less than 25 (P=0.01) or APACHE II 25 or more (P<0.01). The PIRO consistently overestimated actual mortality when stratified by levels of predicted mortality.

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Vorwerk et al (2009),[52] UK **Risk of bias:** Low	To determine the efficacy of the abbreviated Mortality in Emergency Department Sepsis (MEDS) score (without neutrophil bands), the MEWS score and NPT lactate in predicting 28-day mortality in adult ED patients with sepsis.	Retrospective cohort (Validation)	307 adult ED patients (>16 years) with sepsis admitted to 2 hospitals. Patients were excluded if parameters to calculate the MEW or MEDS score were missing. Mean age 69.7 years (95% CI 67.5 to 71.8); 51% men.	**MEWS:** 5 Parameters: SBP, pulse rate, respiratory rate, temperature, AVPU score **Abbreviated MEDS** (Excluded because it is a condition (sepsis) specific score) **Blood lactate** Was only routinely measured in the ED of 1 of the hospitals (n = 158)	28-day mortality	**MEWS** MEWS ≥5 Sensitivity: 72.2% (95% CI 60.4% to 82.1%) Specificity: 59.2% (95% CI 52.6% to 65.5%) Significant predictor for non-survival (OR 3.76; 95% CI 2.11 to 6.71). AUROC: 0.72 (0.67 to 0.77) **Lactate** Lactate level of ≥4 mmol/l Sensitivity: 49.1% (95% CI 35.1% to 63.2%) Specificity: 74.3% (95% CI 64.8% to 82.3%) Significant predictor for non-survival (OR 2.80; 95% CI 1.39 to 5.57). AUROC: 0.62 (0.54 to 0.70)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/tool	Reference criteria	Results
Williams et al. (2016),[77] Australia **Risk of bias:** Low	(1) To validate a number of severity of illness scores in patients admitted with presumed infection; (2) To compare the performance of scores in patient subgroups with increasing mortality: infection without systemic inflammatory response syndrome, sepsis, severe sepsis and septic shock	Prospective cohort (Validation)	8,871 patients admitted with presumed infection, >17 years of age, to ED of 1 metropolitan hospital in 160 weeks in 2 periods: Oct 2007-Dec 2008 (pilot) and June 2009-May 2011 (funded).	**Simplified Acute Physiology Score II (SAPS II)** Parameters: Age, chronic disease, type of admission, GCS, temp, BP, HR, FiO_2/PaO_2/Aa gradient, bicarbonate, sodium, potassium, WBC, bilirubin, urea, urine output. **Sequential Organ Failure Assessment (SOFA)** Parameters: GCS, BP, vasopressor use, FiO_2/PaO_2/Aa gradient, platelets, bilirubin, creatine, urine output. **Acute Physiology and Chronic Health Evaluation II (APACHE II) score** Parameters: Age, chronic disease, type of admission, GCS, temp, BP, HR, FiO2/PaO2/Aa gradient, pH, respiratory rate, sodium, potassium, WBC, haematocrit, creatinine, renal failure. **Severe Sepsis Score (SSS)** (Excluded because it is a condition (sepsis) specific score) **Mortality in ED Sepsis Score (MEDS)** (Excluded because it is a condition (sepsis) specific score)	30-day mortality	30-day mortality: 3.7%. All scores overestimated mortality. Scores in ICU settings overestimated mortality in ED. **AUROC (95% CIs) for Scores, by Sepsis Subgroups** APACHE II Entire Cohort = 0.90 (0.88–0.91) Sepsis: 0.86 (0.84–0.88) Severe Sepsis: 0.79 (0.76–0.83) Septic Shock: 0.79 (0.74–0.84) ICU: 0.77 (0.67–0.87) SAPS II Entire Cohort: 0.90 (0.89–0.92) Sepsis: 0.88 (0.86–0.90) Severe Sepsis: 0.82 (0.78–0.85) Septic Shock: 0.82 (0.78–0.87) ICU: 0.80 (0.71–0.88) SOFA Entire Cohort: 0.86 (0.84–0.88) Sepsis: 0.83 (0.80–0.86) Severe Sepsis: 0.78 (0.75–0.82) Septic Shock: 0.74 (0.68–0.79) ICU: 0.74 (0.65–0.84)

Table 12. Evidence table: Development and validation studies – Undifferentiated patient groups
(Risk of bias was rated using a tool adapted from Kansagara et al (2011).[30] Full details of the risk of bias assessment are available in Appendix 3.)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Burch et al (2008),[45] South Africa **Risk of bias**: High	To evaluate the use of the Modified Early Warning Score (MEWS) as a triage tool to identify medical patients presenting to the emergency department who require admission to hospital and are at increased risk of in-hospital death.	Prospective cohort (Validation)	790 (70.2% of the potential study cohort) medical patients (not surgical, orthopaedic, gynaecological or trauma related) presenting to ED of an urban public hospital.	**MEWS:** Parameters: SBP, pulse rate, respiratory rate, temperature, AVPU score	Admission to hospital, in-hospital mortality	MEWS score median 3 (range 0-11); 26.0% had MEWS score ≥ 5. **Hospital admission** Increased admission with increasing MEWS (trend P-value <0.001). MEWS 0–2 (45%; ref) MEWS 3–4 (59%; RR 1.3; 95% CI 1.1 to 1.6) MEWS ≥5 (79%; RR 1.7; 95% CI 1.5 to 2.0) **In-hospital mortality** Increased mortality with increasing MEWS (trend P-value <0.001). MEWS 0–2 (5%; ref) MEWS 3–4 (16%; RR 2.8; 95% CI 1.7 to 4.8) MEWS ≥5 (26%; RR 4.6; 95% CI 2.7 to 7.8)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Dundar et al (2015),[79] Turkey **Risk of bias:** Low	The aim of this study was to evaluate the value of the Modified Early Warning Score (MEWS) and the VitalPac Early Warning Score (VIEWS) in predicting hospitalization and in-hospital mortality in geriatric emergency department (ED) patients.	Prospective cohort (Validation)	671 (all) patients (aged ≥65 years) presenting to the ED of 1260-bed hospital between 15 January 2014 and 15 February 2014.	**MEWS** Parameters: respiratory rate, SBP, HR, Temp, AVPU **VIEWS** Parameters: respiratory rate, SBP, HR, Temp, oxygen saturation, inhaled oxygen, AVPU	Hospitalization & in-hospital mortality	187 (27.9%) admitted to a ward 153 (22.8%) were admitted to ICU 4 (0.6%) patients died during the follow-up at the ED 8.5% in-hospital mortality rate **Hospitalisation** MEWS (optimal cut-off: 3) AUROC: 0.73 (95% CI 0.69–0.77) Sensitivity: 42%, Specificity: 89% LR+: 3.7, LR−: 0.7 VIEWS (optimal cut-off: 6) AUROC: 0.76 (95% CI 0.72–0.79) Sensitivity: 56%, Specificity: 85% LR+: 3.8, LR−: 0.5. **In-hospital mortality** MEWS (optimal cut-off ≥4) AUROC: 0.89 (95% CI 0.84–0.94) Sensitivity: 74%, Specificity: 89% LR+: 6.7, LR−: 0.3 VIEWS (optimal cut-off: ≥8) AUROC: 0.90 (95% CI 0.86–0.94) Sensitivity: 84%, Specificity: 83% LR+: 4.9, LR−: 0.2. There was no statistical difference between AUROC of MEWS & VIEWS in predicting hospitalization and in-hospital mortality (P=0.28 and 0.82).

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Eick et al (2015),[70] Germany **Risk of bias**: Low	To evaluate heart rate deceleration capacity, an electrocardiogram-based marker of autonomic nervous system activity, as risk predictor in a medical emergency department and to test its incremental predictive value to the Modified Early Warning Score (MEWS).	Prospective cohort (Validation)	5730 consecutive patients ≥18 years admitted to ED of 1 hospital over a 22 month period and in sinus rhythm.	**MEWS** Parameters: respiratory rate, SBP, HR, Temp, LOC at ED admission **Heart rate variability (deceleration capacity)** (excluded because not routine assessment)	**Primary outcome**: intra-hospital mortality **Secondary outcome**: total mortality at 30 and 180 days, transfer to ICU during hospital stay	Admission to ICU (%) n=366 (6.4%) In-hospital deaths: n=142 (2.5%) Deaths at 30 days: n=196 (3.4%) Deaths at 180 days: n=436 (7.6%) Mean (SD) MEWS Survivors 2.3±1.4 vs Non-survivors MEWS= 5.0±1.7 (p< 0.001) **In-hospital mortality:** MEWS: AUROC: 0.71 (0.67–0.75; p< 0.001) Adjusted OR 1.14 (95% CI 1.09–1.19) **Secondary outcomes** not reported for MEWS.
Graham et al (2007),[63] Hong Kong **Risk of bias:** Unclear (Conference abstract only)	To validate the use of a Modified Early Warning Score (MEWS) in ED to identify patients at risk of serious illness requiring hospital admission.	Prospective cohort (Validation)	413 patients (96.5% of eligible patients) admitted to a 16-bed ED observation ward of 1 hospital.	**MEWS** Parameters: respiratory rate, SBP, HR, Temp, LOC	Inpatient hospital admission, reattendance at ED within 48 hrs related to index of diagnosis after discharge, 30 day mortality.	Admitted: 46 Reattended within 48 hrs of discharge (4 required admission): 10 Deceased 30 days: 2 MEWS score >4: • Increased need for hospital admission OR 8.3: 95% CI = 1.1-60.4, p=0.013) • Increased ED reattendance within 48 hrs (OR 45.2: 95% CI=3.4 to 568.9, p< 0.0001) (Limited information; abstract only)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Heitz et al (2010),[54] USA **Risk of bias**: Low	Examine the performance characteristics and discriminatory ability of the most abnormal Modified Early Warning Score (MEWS) (MEWS Max) score during the entire ED stay in predicting the need for higher levels of care among ED patients presenting to a tertiary care facility.	Retrospective cohort (Validation)	280 of 500 randomly selected charts of all patients presenting to the ED of one Medical Center in 2005.	**Adapted MEWS (MEWS Max)**: Parameters: SBP, pulse rate, respiratory rate, temperature, Glasgow Coma Scale (GSC). **MEWS plus**: Parameters: MEWS max, age 60, race, gender, ED length of stay, method of arrival, and antibiotics given prior to or during ED visit.	Composite of: Need for higher level of care (defined as initial admission from the ED or transfer within 24 hours to a non-floor bed (acute care, intermediate care unit, or critical care unit) or mortality within 24 hours of ED presentation.	27% (76/280) met composite outcome of death (n=1) or need for higher care (n=75). **MEWS Max** The MEWS Max was significantly associated with the primary composite outcome ($P < 0.001$, Cochran-Armitage trend test). Optimum threshold MEWS Max: ≥4 Sensitivity: 62% (95% CI 50-73) Specificity: 79% (95% CI 73-84) PPV: 52 NPV: 85 AUROC 0.73 (95% CI, 0.66- 0.79) Each 1-point increase in the MEWS Max score associated with a 60% increase in the odds of meeting the composite endpoint (OR 1.6; 95% CI, 1.3-1.8). **MEWS Plus** AUROC 0.76 (95% CI, 0.69- 0.82) In 58 cases (21.7%), using MEWS Plus would have placed patients in a more appropriate risk category than MEWS Max, while 5.6% of cases would have resulted in inappropriate reclassification.

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Junhasavasdiku et al (2012),[74] Thailand **Risk of bias:** Unclear	To determine whether admission delay (lead-time) and other factors are associated with hospital mortality rates of emergency medical patients. (Modified Early Warning Score (MEWS) data used in this synthesis).	Prospective cohort (Validation)	381 patients, >15 years of age, presenting to ED between Aug-Nov 2009, and admitted to medical wards of a tertiary urban care centre, including intensive care units.	**MEWS:** Parameters: SBP, pulse rate, respiratory rate, temperature, AVPU score	Mortality	Overall mortality rate was 8.9%. MEWS at ED was associated with mortality ($p<0.001$): Non-survivors median 4 (range 1-10), survivors median 2 (range 0-11)
Naidoo et al (2014),[44] South Africa **Risk of bias:** High	To evaluate the use of the Triage Early Warning Score (TEWS) by healthcare workers in an ED in a large urban hospital in KwaZulu-Natal, and its ability to identify patients who require admission and at increased risk for in-hospital mortality.	Retrospective cohort (Validation)	265 patient records in an ED of 1 urban hospital.	**TEWS** : Parameters: Mobility, Resting rate, HR, SBP, AVPU, Trauma	Discharge within 24 hours of admission, admission to a ward, admission to an intensive care unit (ICU), and death in hospital.	47.6% were admitted to wards and 3 (1.1%) admitted to ICU; 4 patients (1.5%) died within 24 hours of admission. 233 (87.9%) had a TEWS < 7, while 32 (12.1%) had a TEWS ≥ 7. A significant association between the TEWS category and outcome was established (no details of significance tests provided): 53.7% of patients with a TEWS of < 7 were discharged, compared to 18.7 % with a score ≥ 7 who were discharged. No patients in the low-score category were admitted to ICU. No patients died. Three patients were admitted to ICU, and four died in the high-score category.

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Olsson et al (2003),[67] Sweden **Risk of bias:** Low	(1) Could the abbreviated severity of disease classification system Rapid Acute Physiology Score (RAPS), created for use in the out-of-hospital setting, be useful in the ED for predicting in-hospital mortality and the length of hospital stay (LOS) in nonsurgical patients? (2) Is it possible to modify RAPS to provide a more potent scoring system (the Rapid Emergency Medicine Score [REMS]) by including age and one or two parameters easily obtained by modern technology (oxygenation and body temperature) for the purpose of predicting in-hospital mortality? (3) Could REMS, with its simplicity and fewer variables, perform as well as APACHE II in the nonsurgical ED?	Prospective cohort (Development & Validation)	1027 adult nonsurgical patients were recruited from two sources: 185 nonsurgical, critically ill patients referred to the ICU from 1 Nov 1995 -1 Nov 1996, and 885 patients at the nonsurgical ED who were admitted either to an ordinary medical department (n = 758), to a general ICU (n = 9), to a coronary care unit (n = 84), or to a neuro-ICU (n = 15) between 1 Jan 1996 -1 March 1996.	**Model Validation** **APACHE II** Parameters: temperature, mean arterial pressure, HR, oxygenation of arterial blood (PaO_2), arterial pH, serum sodium, serum potassium, serum creatinine, haematocrit, white blood cell count, and GCS score. (Arterial pH was not used in the scoring system because this variable is not measured routinely in the ED.) **RAPS** Parameters: HR, BP, respiratory rate, and GCS score. + peripheral oxygen saturation (0–4 points), body temperature (0–4) and age were added to the four RAPS variables. **Model development** **REMS** (based on best predictors of RAPS) Parameters: coma, respiratory frequency, oxygen saturation, BP, and HR (maximal score being 4 for all) and age (maximal score being 6).	In-hospital mortality	Mortality of 116 (11%). **REMS** Likelihood ratio chi-square value of 318.7 (p < 0.0001) OR 1.58 (95% CI 1.48 to 1.70). AUROC: 0.91 +/- 0.02 (had a superior discriminating power compared to RAPS (p<0.001)) **RAPS** Likelihood ratio chi-square value of 273 (p<0.0001) OR 1.77 (95% CI 1.62 to 1.93). AUROC: 0.87 +/- 0.02 **APACHE II** Likelihood ratio chi-square value of 278.5 (p <0.0001) OR 1.25 (95% CI 1.21 to 1.29). AUROC: 0.90 +/- 0.02 (no significant difference with REMS)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Olsson et al (2004),[68] Sweden **Risk of bias:** Low	(1) Could the abbreviated severity of disease classification system Rapid Acute Physiology Score (RAPS), created to be used in the pre-hospital setting, be useful in the ED to predict in-hospital mortality and hospital length of stay (LOS) in nonsurgical patients?; (2) Is it possible to modify RAPS to provide a more powerful scoring system for medical patients (Rapid Emergency Medicine Score, REMS) to predict in-hospital mortality?	Prospective cohort (Development & Validation)	11751 nonsurgical patients presenting to the ED during 12 consecutive months.	**RAPS** Parameters: blood pressure, respiratory rate, pulse rate and Glasgow coma scale Model Developed **REMS** (based on significant predictors of RAPS) Parameters: coma, respiratory frequency, oxygen saturation, blood pressure and pulse rate (maximal score being 4 for all) and age (maximal score being 6)	In-hospital mortality	Mortality: n=285 **RAPS** Likelihood ratio chi-square value of 261.2 (P < 0.0001) OR 1.47 (95% CI: 1.41–1.54). AUROC: 0.65 ± 0.02 **REMS** Likelihood ratio chi-square value of 487.3 (P < 0.0001) OR 1.40 (95% CI: 1.36–1.45) AUROC: 0.85 ± 0.01; superior discriminating power compared with RAPS (P < 0.001)

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Subbe et al (2006),[51] UK **Risk of bias:** Low	To establish a frequency distribution for typical physiological scoring systems and to establish the potential benefit of adding these to an existing triage system in accident and emergency departments.	Retrospective cohort (Validation)	Group 1: 53 unselected patients presenting at ED in two samples of consecutive patients on 30 and 31 Oct 2003. Group 2: 49 direct admissions from ED to the ICU admitted between 1 April-31 Oct 2003. Group 3: 49 patients admitted to ED, who were transferred to a general medical or surgical ward and then admitted to ICU between 1 April-31 Oct 2003. Total of 151 patients.	**MEWS** (Modified Early Warning Score)**:** <u>Parameters:</u> SBP, pulse rate, respiratory rate, temperature, AVPU score **ASSIST** (Assessment Score for Sick patient Identification and Step-up in Treatment) <u>Parameters:</u> SBP, pulse rate, respiratory rate, level of consciousness (ACDN score), age. **MET** (Medical Emergency Team) Criteria for the call-out of a MET based on Airway, Breathing, Circulation, Disability assessment. Nursing staff are asked to call out senior staff if bedside observations are below or above defined thresholds for blood pressure, heart rate, respiratory rate & level of consciousness, or if worried about a patient. <u>Existing triage system</u>**:** **MTS** (Manchester Triage System) Uses protocols based on the presenting complaint and questions about aggravating factors.	Critically ill (defined as MEWS >2, ASSIST >3 and MET criteria applicable, with MTS categories orange or red).	Patients identified as critically ill (at risk of deterioration): **MTS** (orange or red) Group 1: Sensitivity 15% Group 2: Sensitivity 96% Group 3: Sensitivity 65% **MEWS** (>2) Group 1: Sensitivity 8% Group 2: Sensitivity 77% Group 3: Sensitivity 55% **ASSIST** (>3) Group 1: Sensitivity 0% Group 2: Sensitivity 22% Group 3: Sensitivity 16% **MET** (=1) Group 1: 0 Group 2: Sensitivity 2% Group 3: Sensitivity 7%

Authors (year), country	Study aim	Study design (type)	Participants	Content of system/ tool	Reference criteria	Results
Wang et al (2016),[75] Taiwan **Risk of bias:** Unclear	To evaluate whether peri-arrest Modified early Warning Score (MEWS) could be a prognostic factor in in-hospital cardiac arrest (IHCA). To combine pre-arrest comorbidity factors (Charleson Comorbidity Index, CCI), peri-arrest physiological factors (MEWS) and arrest factors to evaluate the outcome of IHCA in ED.	Retrospective cohort (Validation)	99 non-traumatic, >20 years of age, patients ED records of one hospital over 30 months period.	**Charlson Comorbidity Index** 1-6 points allocated by morbidity. **Peri-arrest MEWS** Temp, BP, HR, Respiratory Rate, LoC (AVPU) from triage to 0.5hrs prior to arrest (peri-arrest MEWS)	In-hospital cardiac arrest, Survival to discharge (STDG)	Lower CCI in STDG group (n=22) than mortality group (n=77) (2.27±1.87 vs 3.87±2.83; p=0.001) No significant difference in MEWS at triage between STDG vs mortality group (3.42± 2.2 vs 4.02 ± 2.65; p=0.81) Lower periarrest MEWS in STDG vs mortality group (4.41± 2.28 vs 5.82 ± 2.84; p=0.05) **Survival to discharge** CCI Adjusted OR 0.57 (95% CI 0.38-0.84); p=0.005 Peri-arrest MEWS Adjusted OR 0.77 (95% CI 0.60-0.97); p=0.028

4.3.6 Health economics

We did not identify any formal evaluations that examine the cost effectiveness of early warning systems or TTS or other scoring systems in hospital Emergency Departments. While it is clear that implementing early warning systems requires a healthcare resource investment, the degree to which such systems may or may not result in cost savings elsewhere in the healthcare system, or in improved patient outcomes, remains unclear. As described earlier in this report, the limited evidence base suggests that early warning systems are effective in, for example, identifying deteriorating patients, reducing cardiac arrests and reducing intensive care unit admissions. Should these effects exist, the potential healthcare cost savings could go to fund, at least to some degree, the implementation of early warning systems in ED clinical practice. While this theory is open to question, it does go to highlight the need for primary research studies to be conducted to directly evaluate the cost effectiveness of early warning systems. Such studies should focus on the monitoring of resource use, costs and patient outcomes in order to determine whether early warning systems are likely to deliver good value for money.

Conclusion

Implications for practice

Five objectives were addressed in this review. The first objective was to describe the use of early warning systems in the ED. Multiple early warning systems were identified but the extent to which they are used in the ED seems to vary in different countries from which data was available (UK and Australia). Ten descriptive studies included in this review demonstrated that the use of early warning systems in ED was linked with an increase in escalation protocol activation, but incorrect calculation of scores was common. Compliance with recording early warning system scores was relatively low, although the vital signs HR and BP were usually recorded. This finding emphasises the importance of effective implementation strategies. However, we did not identify any studies examining educational programmes for early warning systems (objective 5). Existing guidelines regarding the use of early warning systems to monitor acute patients in hospital did include educational tools, but were not specific to the ED. The three guidelines we identified all recommend inclusion of the following six parameters: respiratory rate, heart rate, systolic blood pressure, temperature, oxygen saturations, and level of consciousness.

Evidence from 35 validation and development studies, assessing 27 different systems, demonstrated that early warning systems used in ED settings seem to be able to predict adverse outcomes including mortality, admission to hospital or ICU, and length of hospital stay, but there is variability between studies (objective 3). All but two early warning systems were aggregated scores. This limited the ability to compare comprehensively between single, multiple parameter and aggregated scores. The APACHE II score, PEDS, VIEWS-L, and THERM scores were relatively best at predicting mortality and ICU admission, providing excellent discrimination ability (AUROC > 0.8),[83] but differences between studies may, in part, account for this. The MEWS was the most commonly used and assessed system, but findings of this review suggest a relatively lower ability to predict mortality and ICU admissions compared to the four scores mentioned above, with only some studies indicating acceptable discriminatory ability of the MEWS (AUROC > 0.7) and other studies indicating a lack of discriminatory ability (AUROC < 0.7),[83] especially for the outcome ICU admission. The exception was one study that found excellent discriminatory ability of MEWS for the outcome in-hospital mortality (AUROC 0.89).[79] However, the ability of early warning systems to predict adverse outcomes does not mean that early warning systems are effective at preventing adverse outcomes. Only one study was identified that addressed this question and it found that the introduction of an early warning system may have little or no difference in detecting deterioration or adverse events however the evidence was of very low quality making it impossible to draw any strong conclusions (objective 2). In addition, we did not find any studies examining the cost-effectiveness of early warning systems and TTS (objective 4).

Implications for research

There is a clear need for high quality effectiveness studies to test the impact of using early warning systems or TTS in the ED on patient outcomes. The cost-effectiveness of such interventions, the effectiveness of related educational programmes, and barriers and facilitators to implementation also need to be examined, as there is a clear lack of such evidence.

Limitations

This was a rapid review; however, current literature suggests that putting some restrictions in place in the review process does not significantly impact on the quality and conclusions of a review.[84] The restrictions for this review included no translation of reports, although only one non-English study was identified. Data extraction was initially done by only one reviewer, but 50% of records where checked for accuracy by a second reviewer. Strengths of the review lie in its thorough search strategy, its scope and inclusion of different studies and reports to address the research objectives, and in its rigorous methodology with dual independent screening and quality assessment.

References

1. Goldhill DR, McNarry AF. Physiological abnormalities in early warning scores are related to mortality in adult inpatients. British journal of anaesthesia. 2004;92(6):882-4.
2. Buist M, Bernard S, Nguyen TV, Moore G, Anderson J. Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study. Resuscitation. 2004;62(2):137-41.
3. Schein RM, Hazday N, Pena M, Ruben BH, Sprung CL. Clinical antecedents to in-hospital cardiopulmonary arrest. Chest. 1990;98(6):1388-92.
4. Smith AF, Wood J. Can some in-hospital cardio-respiratory arrests be prevented? A prospective survey. Resuscitation. 1998;37(3):133-7.
5. Franklin C, Mathew J. Developing strategies to prevent inhospital cardiac arrest: analyzing responses of physicians and nurses in the hours before the event. Critical care medicine. 1994;22(2):244-7.
6. Rivers EP, Coba V, Whitmill M. Early goal-directed therapy in severe sepsis and septic shock: a contemporary review of the literature. Current opinion in anaesthesiology. 2008;21(2):128-40.
7. Department of Health. National Early Warning Score: National Clinical Guideline No.1. Dublin: Department of Health, 2013.
8. Royal College of Physicians. National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. Report of a working party. London: Royal College of Physicians, 2012.
9. National Institute for Health and Care Excellence. Acutely Ill Patients in Hospital. Recognition of and Response to Acute Illness in Adults in Hospital (CG50). London: NICE, 2007.
10. Mitchell IA, McKay H, Van Leuvan C, Berry R, McCutcheon C, Avard B, et al. A prospective controlled trial of the effect of a multi-faceted intervention on early recognition and intervention in deteriorating hospital patients. Resuscitation. 2010;81(6):658-66.
11. Agency for Healthcare Research and Quality. Rapid Response Systems USA: AHRQ, U.S. Department of Health and Human Services, USA; 2014 [Available from: https://psnet.ahrq.gov/primers/primer/4/rapid-response-systems.
12. Department of Health. The Irish Maternity Early Warning System (IMEWS): National Clinical Guideline No. 4. Dublin: Department of Health, 2014.
13. Singh S, McGlennan A, England A, Simons R. A validation study of the CEMACH recommended modified early obstetric warning system (MEOWS)*. Anaesthesia. 2012;67(1):12-8.
14. Maguire P, O'Higgins A, Power K, Turner M. The Irish Maternity Early Warning System (IMEWS). Irish Medical Journal. 2014;107(10):309.
15. Department of Health. The Irish Paediatric National Early Warning System (PEWS): National Clinical Guideline No. 12. Dublin: Department of Health, 2015.
16. Smith M, Feied C. The Emergency Department as a Complex System 1999 [Available from: http://www.necsi.edu/projects/yaneer/emergencydeptcx.pdf.
17. Courtney DM, Neumar RW, Venkatesh AK, Kaji AH, Cairns CB, Lavonas E, et al. Unique characteristics of emergency care research: scope, populations, and infrastructure. Academic emergency medicine: official journal of the Society for Academic Emergency Medicine. 2009;16(10):990-4.
18. Roland D, Coats TJ. An early warning? Universal risk scoring in emergency medicine. Emergency Medicine Journal. 2011;28(4):263.
19. Griffiths JR, Kidney EM. Current use of early warning scores in UK emergency departments. Emergency Medicine Journal. 2012;29(1):65-6 2p.
20. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified early warning score in medical admissions. QJM - Monthly Journal of the Association of Physicians. 2001;94(10):521-6.
21. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ (Clinical research ed). 2009;339.
22. The Nordic Cochrane Centre; The Cochrane Collaboration. Review Manager (RevMan) [Computer program] Version 5.3. Copenhagen 2014.

23. National Heart Lung and Blood Institute. Study Quality Assessment Tools. Secondary Study Quality Assessment Tools 2014 [Available from: http://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools.
24. The AGREE collaboration GW, Cluzeau FA, et al,. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Quality and Safety in Health Care. 2003;12(1):18-23.
25. Higgins J, Green S. Cochrane Handbook for systematic Reviews of Interventions Version 5.1.0 [updated March 2011]: The Cochrane Collaboration; 2011. Available from: www.cochrane-handbook.org.
26. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ (Clinical research ed). 2004;328(7454):1490.
27. Effective Practice and Organisation of Care (EPOC). Suggested risk of bias criteria for EPOC reviews. EPOC Resources for review authors Oslo: Norwegian Knowledge Centre for the Health Services; 2015 [Available from: http://epoc.cochrane.org/epoc-specific-resources-review-authors]
28. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ : British Medical Journal. 1996;313(7052):275-83.
29. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol Assess. 2004;8(36):iii-iv, ix-xi, 1-158.
30. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: A systematic review. JAMA. 2011;306(15):1688-98.
31. Abrami P, Borokhovski E, Bernard R, Wade A, Tamim R, Persson T, et al. Issues in conducting and disseminating brief reviews of evidence. Evidence & Policy. 2010;6(3):371-89
32. Gu M, Fu Y, Li C, Chen M, Zhang X, Xu J, et al. The value of modified early warning score in predicting early mortality of critically ill patients admitted to emergency department. Chinese Critical Care Medicine. 2015;27(8):687-90.
33. Coughlan E, Deasy C, McDaid F, Geary U, Ward M, O'Sullivan R, et al. An introduction to the Emergency Department Adult Clinical Escalation Protocol: ED-ACE. Resuscitation. 2015;96:33- 1p.
34. Considine J, Lucas E, Wunderlich B. The uptake of an early warning system in an Australian emergency department: A pilot study. Critical Care and Resuscitation. 2012;14(2):135-41.
35. Correia N, Rodrigues RP, Sá MC, Dias P, Lopes L, Paiva A. Improving recognition of patients at risk in a Portuguese general hospital: results from a preliminary study on the early warning score. International Journal of Emergency Medicine. 2014;7(1):Article number 22.
36. Griffiths JR, Kidney EM. Current use of early warning scores in UK emergency departments. Emergency medicine journal : EMJ. 2012;29(1):65-6.
37. Wilson SJ, Wong D, Clifton D, Fleming S, Way R, Pullinger R, et al. Track and trigger in an emergency department: an observational evaluation study. Emergency Medicine Journal. 2013;30(3):186-91 6p.
38. Christensen D, Jensen NM, Maaloe R, Rudolph SS, Belhage B, Perrild H. Low compliance with a validated system for emergency department triage. Danish Medical Bulletin. 2011;58(6):A4294.
39. Johnson KD, Winkelman C, Burant CJ, Dolansky M, Totten V. The Factors that Affect the Frequency of Vital Sign Monitoring in the Emergency Department. JEN: Journal of Emergency Nursing. 2014;40(1):27-35 9p.
40. Austen C, Patterson C, Poots A, Green S, Weldring T, Bell D. Using a local early warning scoring system as a model for the introduction of a national system. Acute Medicine. 2012;11(2):66-73.
41. Hudson P, Ekholm J, Johnson M, Langdon R. Early identification and management of the unstable adult patient in the emergency department. Journal of Clinical Nursing. 2015;24(21-22):3138-46.
42. Shuk-Ngor S, Chi-Wai O, Lai-Yee W, Chung JYM, Graham CA. Is the Modified Early Warning Score able to enhance clinical observation to detect deteriorating patients earlier in an Accident & Emergency Department? Australasian Emergency Nursing Journal. 2015;18(1):24-32 9p.

43. Wilson SJ, Wong D, Pullinger RM, Way R, Clifton DA, Tarassenko L. Analysis of a data-fusion system for continuous vital sign monitoring in an emergency department. European Journal of Emergency Medicine. 2016;23(1):28-32.
44. Naidoo DK, Rangiah S, Naidoo SS. An evaluation of the Triage Early Warning Score in an urban accident and emergency department in KwaZulu-Natal. South African Family Practice. 2014;56(1):69-73 5p.
45. Burch VC, Tarr G, Morroni C. Modified early warning score predicts the need for hospital admission and inhospital mortality. Emergency Medicine Journal. 2008;25(10):674-8 5p.
46. Challen K, Goodacre SW. Predictive scoring in non-trauma emergency patients: a scoping review. Emergency Medicine Journal. 2011;28(10):827-37 11p.
47. Australian Commission on Safety and Quality in Health Care (the Commission). Recognising and responding to clinical deterioration, Survey of recognition and response systems in australia. 2011.
48. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters. 2006;27: 861–74.
49. Corfield AR, Lees F, Houston G, Zealley I, Dickie S, Ward K, et al. Early warning scores in sepsis: Utility of a single early warning score in the emergency department? Intensive Care Medicine. 2012;38:S296-S7.
50. Keep JW, Messmer AS, Sladden R, Burrell N, Pinate R, Tunnicliff M, et al. National early warning score at Emergency Department triage may allow earlier identification of patients with severe sepsis and septic shock: A retrospective observational study. Emergency Medicine Journal. 2016;33(1):37-41.
51. Subbe CP, Slater A, Menon D, Gemmell L. Validation of physiological scoring systems in the accident and emergency department. Emergency Medicine Journal. 2006:841-5 5p.
52. Vorwerk C, Loryman B, Coats TJ, Stephenson JA, Gray LD, Reddy G, et al. Prediction of mortality in adult emergency department patients with sepsis. Emergency Medicine Journal. 2009;26(4):254-8 5p.
53. Albright CM, Ali TN, Lopes V, Rouse DJ, Anderson BL. The Sepsis in Obstetrics Score: a model to identify risk of morbidity from sepsis in pregnancy. American Journal of Obstetrics & Gynecology. 2014;211(1):39.e1-8 1p.
54. Heitz CR, Gaillard JP, Blumstein H, Case D, Messick C, Miller CD. Performance of the maximum modified early warning score to predict the need for higher care utilization among admitted emergency department patients. Journal of Hospital Medicine. 2010;5(1):E46-52 1p.
55. Howell MD, Donnino MW, Talmor D, Clardy P, Ngo L, Shapiro NI. Performance of severity of illness scoring systems in emergency department patients with infection. Academic Emergency Medicine. 2007;14(8):709-14 6p.
56. Jones AE, Fitch MT, Kline JA. Operational performance of validated physiologic scoring systems for predicting in-hospital mortality among critically ill emergency department patients. Critical care medicine. 2005;33(5):974-8 5p.
57. Nguyen HB, Van Ginkel C, Batech M, Banta J, Corbett SW. Comparison of Predisposition, Insult/ Infection, Response, and Organ dysfunction, Acute Physiology And Chronic Health Evaluation II, and Mortality in Emergency Department Sepsis in patients meeting criteria for early goal-directed therapy and the. Journal of Critical Care. 2012;27(4):362-9 8p.
58. Armagan E, Yilmaz Y, Olmez OF, Simsek G, Gul CB. Predictive value of the modified early warning score in a turkish emergency department. European Journal of Emergency Medicine. 2008;15(6):338-40.
59. Bulut M, Cebicci H, Sigirli D, Sak A, Durmus O, Top AA, et al. The comparison of modified early warning score with rapid emergency medicine score: A prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. Emergency medicine journal [Internet]. 2014; 31(6):[476-81 pp.]. Available from: http://emj.bmj.com/content/31/6/476.full.pdf.
60. Çildir E, Bulut M, Akalin H, Kocabaş E, Ocakoğlu G, Aydin ŞA. Evaluation of the modified MEDS, MEWS score and Charlson comorbidity index in patients with community acquired sepsis in the emergency department. Internal and Emergency Medicine. 2013;8(3):255-60.
61. Cattermole GN, Mak SKP, Liow CHE, Ho MF, Hung KYG, Keung KM, et al. Derivation of a prognostic score for identifying critically ill patients in an emergency department resuscitation room. Resuscitation. 2009;80(9):1000-5.

62. Cattermole GN, Liow ECH, Graham CA, Rainer TH. THERM: the Resuscitation Management score. A prognostic tool to identify critically ill patients in the emergency department. Emergency Medicine Journal. 2014;31(10):803-7 5p.
63. Graham CA, Choi KM, Ki CW, Leung YK, Leung PH, Mak P, et al. Evaluation and validation of the use of Modified Early Warning Score (MEWS) in emergency department observation ward...2007 Society for Academic Emergency Medicine Annual Meeting. Academic Emergency Medicine. 2007;14:S199-200 1p.
64. Ho LO, Li H, Shahidah N, Koh ZX, Sultana P, Ong MEH. Poor performance of the modified early warning score for predicting mortality in critically ill patients presenting to an emergency department. World Journal of Emergency Medicine. 2013;4(4):273-7.
65. Hock Ong ME, Lee Ng CH, Goh K, Liu N, Koh ZX, Shahidah N, et al. Prediction of cardiac arrest in critically ill patients presenting to the emergency department using a machine learning score incorporating heart rate variability compared with the modified early warning score. Critical Care. 2012;16(3).
66. Liu FY, Qin J, Wang RX, Fan XL, Wang J, Sun CY, et al. A prospective validation of national early warning score in emergency intensive care unit patients at Beijing. Hong Kong Journal of Emergency Medicine. 2015;22(3):137-44.
67. Olsson T, Lind L. Comparison of the Rapid Emergency Medicine Score and APACHE II in nonsurgical emergency department patients. Academic Emergency Medicine. 2003;10(10):1040-8 9p.
68. Olsson T, Terent A, Lind L. Rapid Emergency Medicine score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients. Journal of Internal Medicine. 2004;255(5):579-87 9p.
69. Christensen D, Jensen NM, Maaløe R, Rudolph SS, Belhage B, Perrild H. Nurse-administered early warning score system can be used for emergency department triage. Danish Medical Bulletin. 2011;58(6).
70. Eick C, Rizas KD, Meyer-Zürn CS, Groga-Bada P, Hamm W, Kreth F, et al. Autonomic nervous system activity as risk predictor in the medical emergency department: a prospective cohort study. Critical care medicine. 2015;43(5):1079-86 8p.
71. Geier F, Popp S, Greve Y, Achterberg A, Glockner E, Ziegler R, et al. Severity illness scoring systems for early identification and prediction of in-hospital mortality in patients with suspected sepsis presenting to the emergency department. Wiener klinische Wochenschrift. 2013;125(17-18):508-15.
72. Jo S, Jeong T, Lee JB, Jin Y, Yoon J, Park B. Validation of modified early warning score using serum lactate level in community-acquired pneumonia patients. The National Early Warning Score-Lactate score. American Journal of Emergency Medicine. 2016.
73. Jo S, Lee JB, Jin YH, Jeong TO, Yoon JC, Jun YK, et al. Modified early warning score with rapid lactate level in critically ill medical patients: the ViEWS-L score. Emergency Medicine Journal. 2013;30(2):123-9 7p.
74. Junhasavasdikul D, Theerawit P, Kiatboonsri S. Association between admission delay and adverse outcome of emergency medical patients. Emergency Medicine Journal. 2013;30(4):320-3 4p.
75. Wang AY, Fang CC, Chen SC, Tsai SH, Kao WF. Periarrest Modified Early Warning Score (MEWS) predicts the outcome of in-hospital cardiac arrest. Journal of the Formosan Medical Association. 2016;115(2):76-82 7p.
76. Alam N, Vegting IL, Houben E, van Berkel B, Vaughan L, Kramer MH, et al. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. Resuscitation. 2015;90:111-5.
77. Williams J, Greenslade J, Chu K, Brown A, Lipman J. Severity Scores in Emergency Department Patients With Presumed Infection: A Prospective Validation Study. Critical care medicine. 2016;44(3):539-47 9p.
78. Considine J, Rawet J, Currey J. The effect of a staged, emergency department specific rapid response system on reporting of clinical deterioration. Australasian Emergency Nursing Journal. 2015;18(4):218-26.

79. Dundar ZD, Ergin M, Karamercan MA, Ayranci K, Colak T, Tuncar A, et al. Modified Early Warning Score and VitalPac Early Warning Score in geriatric patients admitted to emergency department. European journal of emergency medicine : official journal of the European Society for Emergency Medicine. 2015.
80. Churpek MM, Yuen TC, Edelson DP. Risk stratification of hospitalized patients on the wards. Chest. 2013;143(6):1758-65.
81. Corfield AR, Lees F, Zealley I, Houston G, Dickie S, Ward K, et al. Utility of a single early warning score in patients with sepsis in the emergency department. Emergency medicine journal : EMJ. 2014;31(6):482-7.
82. Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden Index and Optimal Cut-Point Estimated from Observations Affected by a Lower Limit of Detection. Biometrical journal Biometrische Zeitschrift. 2008;50(3):419-30.
83. Hosmer D, Lemeshow S. Applied Logistic Regression. 2 ed. New York: John Wiley & Sons, Inc. ; 2000.
84. Watt A, Cameron A, Sturm L, Lathlean T, Babidge W, Blamey S, et al. Rapid versus full systematic reviews: validity in clinical practice? ANZ J Surg. 2008;78(11):1037-40.

Appendix 1: Search strategies

Cochrane Library Search Strategy		
Platform: John Wiley and Sons		
Date of search: 04/03/2016		
ID	**Search**	**Hits**
#1	MeSH descriptor: [Emergency Service, Hospital] explode all trees	2078
#2	Emergency near/1 (care or ward or wards or department or departments or unit or units or room or room or medic* or health or healthcare or hospital or service or services or centre or centres or center or centers or treatment or treatments or outpatient or outpatients)	12379
#3	Casualty near/2 (care or ward or wards or department or departments or unit or units or room or room or service or services or centre or centres or center or centers or outpatient or outpatients)	68
#4	Trauma near/2 (care or ward or wards or department or departments or unit or units or room or room or medic* or hospital or service or services or centre or centres or center or centers or outpatient or outpatients)	1505
#5	Triage near/2 (care or ward or wards or department or departments or unit or units or room or room or medic* or hospital or service or services or centre or centres or center or centers or outpatient or outpatients)	98
#6	“accident and emergency” or “accident & emergency” or A&E or “A & E” or “A and E”	1431
#7	{or #1-#6}	14493
#8	Warning near/2 (early or system or systems or score or scores)	168
#9	Trigger near/2 track	6
#10	Trigger near/4 (score or scores or scoring)	20
#11	Escalation near/2 (protocol or protocols or policy or policies or procedure or procedures or guideline or guidelines or guidance)	25
#12	EWS or MEWS	75
#13	POTTS	177
#14	{or #8-#13}	443
#15	adult or adults or adulthood	402614
#16	#7 and #14 and #15	38

Medline Search Strategy		
Platform: Ovid MEDLINE In-Process & Other Non-Indexed Citations and Ovid MEDLINE 1946 to Present		
Date of Search: 04/03/2016		
ID	**Search**	**Hits**
#1	emergency service, hospital.sh.	48807
#2	(Emergency adj (care or ward or wards or department or departments or unit or units or room or room or medic* or health or healthcare or hospital or service or services or centre or centres or center or centers or treatment or treatments or outpatient or outpatients)).af.	197858
#3	(Casualty adj2 (care or ward or wards or department or departments or unit or units or room or room or service or services or centre or centres or center or centers or outpatient or outpatients)).af.	1309
#4	(Trauma adj2 (care or ward or wards or department or departments or unit or units or room or room or medic* or hospital or service or services or centre or centres or center or centers or outpatient or outpatients)).af.	47447
#5	(Triage adj2 (care or ward or wards or department or departments or unit or units or room or room or medic* or hospital or service or services or centre or centres or center or centers or outpatient or outpatients)).af.	1679
#6	("accident and emergency" or "accident & emergency" or A&E or "A & E" or "A and E").af.	1886693
#7	1 or 2 or 3 or 4 or 5 or 6	2093553
#8	(Warning adj2 (early or system or systems or score or scores)).af.	4296
#9	(Trigger adj2 track).af.	67
#10	(Trigger adj4 (score or scores or scoring)).af.	56
#11	(Escalation adj2 (protocol or protocols or policy or policies or procedure or procedures or guideline or guidelines or guidance)).af.	168
#12	(EWS or MEWS).af.	2180
#13	POTTS.af.	5046
#14	8 or 9 or 10 or 11 or 12 or 13	11541
#15	(adult or adults or adulthood).af.	4789831
#16	7 and 14 and 15	362

Embase Search Strategy		
Platform: Elsevier		
Date of Search: 22/02/2016		
ID	**Search**	**Hits**
#1	'emergency ward'/exp	82019
#2	emergency NEAR/1 (care OR ward OR wards OR department OR departments OR unit OR units OR room OR rooms OR medic* OR health OR healthcare OR hospital OR service OR services OR center OR centers OR centre OR centres OR treatment OR treatments OR outpatient OR outpatients)	317266
#3	casualty NEAR/2 (care OR ward OR wards OR department OR departments OR unit OR units OR room OR room OR service OR services OR centre OR centres OR center OR centers OR outpatient OR outpatients)	1741
#4	trauma NEAR/2 (care OR ward OR wards OR department OR departments OR unit OR units OR room OR room OR medic* OR hospital OR service OR services OR centre OR centres OR center OR centers OR outpatient OR outpatients)	51937
#5	triage NEAR/2 (care OR ward OR wards OR department OR departments OR unit OR units OR room OR room OR medic* OR hospital OR service OR services OR centre OR centres OR center OR centers OR outpatient OR outpatients)	1827
#6	accident and emergency' OR 'accident & emergency' OR a&e OR 'a & e' OR 'a and e'	168515
#7	#1 OR #2 OR #3 OR #4 OR #5 OR #6	512537
#8	warning NEAR/2 (early OR system OR systems OR score OR scores)	5852
#9	trigger NEAR/2 track	110
#10	trigger NEAR/4 (score OR scores OR scoring)	116
#11	escalation NEAR/2 (protocol OR protocols OR policy OR policies OR procedure OR procedures OR guideline OR guidelines OR guidance)	341
#12	ews OR mews	4360
#13	potts	4486
#14	#8 OR #9 OR #10 OR #11 OR #12 OR #13	14723
#15	adult OR adults OR 'adulthood'	5624653
#16	#7 AND #14 AND #15	254

CINAHL Complete Search Strategy		
Platform: EBSCOhost		
Date of Search: 04/03/2016		
ID	**Search**	**Hits**
S1	(MH "Emergency Service")	34352
S2	Emergency N1 (care OR ward or wards OR department or departments OR unit or units OR room or rooms OR medic* OR health OR healthcare OR hospital OR service or services OR center or centers OR centre or centres OR treatment or treatments OR outpatient or outpatients)	99485
S3	Casualty N2 (care OR ward or wards OR department or departments OR unit or units OR room or rooms OR service or services OR center or centers OR centre or centres OR outpatient or outpatients)	402
S4	Trauma N2 (care or ward or wards or department or departments or unit or units or room or room or medic* or hospital or service or services or centre or centres or center or centers or outpatient or outpatients)	9842
S5	Triage N2 (care or ward or wards or department or departments or unit or units or room or room or medic* or hospital or service or services or centre or centres or center or centers or outpatient or outpatients)	978
S6	"accident and emergency" or "accident & emergency" or A&E or "A & E" or "A and E"	52174
S7	S1 OR S2 OR S3 OR S5 OR S6	149039
S8	Warning N2 (early or system or systems or score or scores)	1131
S9	Trigger N2 track	44
S10	Trigger N4 (score or scores or scoring)	26006
S11	Escalation N2 (protocol or protocols or policy or policies or procedure or procedures or guideline or guidelines or guidance)	37
S12	ews OR mews	196
S13	potts	184
S14	S8 OR S9 OR S10 OR S11 OR S12 OR S13	27422
S15	adult OR adults OR adulthood	892275
S16	S7 AND S14 AND S15	653

Cost Effectiveness Resources			
Date of Search: 11/03/2016 (except * searched 04/03/2016)			
Website/Database	**URL**	**Search Terms**	**Hits**
Health Technology Assessment Database, NHS Economic Evaluation Database (NHSEED) & Health Economic Evaluation Database (HEED) via The Cochrane Library*	www.cochranelibrary.com	See Cochrane Library tab	n/a
NHS Service Delivery and Organisation (SDO) Research and Development Programme	www.nets.nihr.ac.uk/programmes/hsdr	patient deterioration emergency department	0
		patient deterioration emergency	4. After sifting = 0
		patient deterioration	46. After sifting = 0
		early warning	13. After sifting = 1
		track and trigger	5. After sifting = 0
National Coordinating Centre for Health Technology Assessment (NCCHTA)	www.nets.nihr.ac.uk/programmes/hta	patient deterioration emergency department	0
		patient deterioration emergency	4. After sifting = 0
		patient deterioration	46. After sifting = 0
		early warning	13. After sifting = 1. Same as result for website above, hence discarded
		track and trigger	5. After sifting = 0
NIHR-HTA Database	http://www.crd.york.ac.uk/CRDWeb/	patient deterioration. Filters: HTA published and HTA in progress	2

Guidance Resources			
Date of Search: 13/03/2016 (except * searched 11/03/2016)			
Website/ Database	**URL**	**Search Terms**	**Hits**
Department of Health (including National Clinical Guidelines)*	via Google Advanced Search https://www.google.com/advanced_search	emergency adult* warning OR OR OR triage OR OR OR care OR OR OR trauma OR OR OR trigger OR OR OR esclat* OR OR OR EWS OR OR OR MEWS OR OR OR POTTS . Site filter:health.gov.ie Region filter: Ireland	495. After sifting first 200 hits = 4 . Plus 2 added from brief manual search.
Health Service Executive (HSE)*	via Google Advanced Search https://www.google.com/advanced_search	emergency adult* warning OR OR OR triage OR OR OR care OR OR OR trauma OR OR OR trigger OR OR OR esclat* OR OR OR EWS OR OR OR MEWS OR OR OR POTTS. Site filter:hse.ie Region filter: Ireland	1880. After sifting first 200 hits = 4.
Health Information and Quality Authority (HIQA)	www.hiqa.ie	“emergency department patient deterioration” in keyword box	92. After sifting = 0
National Institute for Health and Care Excellence (NICE)	https://www.nice.org.uk/guidance	patient deterioration emergency department	55. After sifting = 4
		“early warning” or “track and trigger” or ews or mews or potts	38. After sifting = 5
NHS Evidence (incorporating Scottish Intercollegiate Guidelines Network (SIGN) & Guidelines International Network (GIN))	https://www.evidence.nhs.uk	"patient deterioration" and "emergency department" and ("early warning" or "track and trigger" or ews or mews or potts)	17. After sifting = 1
		emergency department and ("early warning" or "track and trigger" or ews or mews or potts)	419. After sifting = 3
Agency for Healthcare Research and Quality (AHRQ) National Guideline Clearinghouse	http://www.guideline.gov	(triage or casualty or trauma or "emergency department")' and '("early warning" or trigger or ews or mews or potts)' Filters: Adult (19 to 44 years) Aged (65 to 79 years) Aged, 80 and over	76. after sifting = 0
		Patient deterioration	187. After sifting = 0

Professional Bodies			
Date of Search: 09-11/03/2016			
Website/ Database	**URL**	**Search Terms**	**Hits**
Irish Association for Emergency Medicine (IAEM)	www.iaem.ie	Single search box, no instructions. Boolean operators are accepted but only with certain search terms the more options added, the less it seems to work. Therefore, a manual search was performed.	0
Royal College for Emergency Medicine (UK)	www.rcem.ac.uk	Single search box, no instructions. Boolean operators are accepted but only with certain search terms the more options added, the less it seems to work. Therefore, a manual search was performed.	3
European Society for Emergency Medicine (EuSEM)	www.eusem.org	No results for simple searches such as: “patient deterioration”, “track and trigger” or “early warning”.	0
American Academy of Emergency Medicine (AAEM)	www.aaem.org	(emergency OR trauma OR casualty OR triage) AND (care OR ward OR wards OR department OR departments OR unit OR units OR room OR rooms OR health OR healthcare OR hospital OR service) AND (warning OR trigger OR EWS OR MEWS OR POTTS) AND (deteriorate OR deterioration OR deteriorated OR deteriorates OR worse OR worsen OR worsening OR adverse OR weaken OR weakened OR weakens OR weaker OR “acute illness”) AND (Monitor or monitors or monitored or monitoring OR escalate OR escalates OR escalated OR escalation OR escalating OR reassess OR reassesses or reassessed or reassessment OR reassessing) AND (adult OR adults OR adulthood)	67 Hits. After sifting = 0

American College of Emergency Physicians (ACEP)	www.acep.org	Single search box, no instructions. Boolean operators not accepted.	1
Society for Academic Emergency Medicine (SAEM)	www.saem.org	(emergency OR trauma OR casualty OR triage) AND (care OR ward* OR department* OR unit* OR room* OR health* OR hospital OR service) AND (warning OR trigger OR EWS OR MEWS OR POTTS) AND (deteriorat* OR worse* OR adverse OR weaken* OR weaker OR "acute illness") (Monitor* OR escalate* OR reassess*) AND (adult*)	4. After sifting = 0
Canadian Association of Emergency Physicians	www.caep.ca	(emergency OR trauma OR casualty OR triage) AND (escalate* OR trigger OR warning OR ews OR potts)	110. After sifting = 0
Australasian Society for Emergency Medicine (ASEM)	www.asem.org.au	Manual[1]	0
Australasian College of Emergency Medicine (ACEM)	www.acem.org.au	Manual	1
International Federation for Emergency Medicine (IFEM)	http://www.ifem.cc	Manual	2
Faculty of Emergency Nursing	www.fen.uk.com	Manual	0
RCN Emergency Care Association	www.rcn.org.uk	Manual	0
Emergency Nurses Association	www.ena.org	Manual	0
Canadian Emergency Nurses	www.nena.ca	Access to this site not allowed for security reasons	
European Society for Emergency Nursing	www.eusen.org	Manual	0
Emergency Nursing weblinks	www.enw.org	(emergency OR trauma OR casualty OR triage) AND (escalate OR trigger OR warning OR ews OR potts)	5. After sifting = 0

Philippine Society of Emergency Care Nurses	www.philippine nursingdirectory.com/associations/philippine-society-of-emergency-care-nurses-psecn/	Manual	3. After sifting = 0
Emergency Nursing Society of South Africa	http://emssa.org.za/enssa/	Manual	0
Australian College of Emergency Nursing	www.acen.com.au	Manual	0
College of Emergency Nursing Australia	www.cena.org.au	Manual	0
College of Emergency Nurses (New Zealand) CENNZ - NZNO	www.nzno.org.nz/colleges/college_of_emergency_nurses	Manual	0
Hong Kong Emergency Nursing	www.hkena.org	Manual	0

[1]Manual searches were performed on websites where regular electronic searching attempts were not useful. This is caused by reduced searching functionality such as search boxes that can only search one word, or a lack of search box. Manual searching involves exploring clickable webpage content e.g., tabs, buttons, hyperlinks etc. in an iterative way to identify relevant resources.

Grey Literature			
Date of Search: 12-13/03/2016			
Website/Database	**URL**	**Search Terms**	**Hits**
RIAN	http://rian.ie/en/static/User#search	All of 'patient', 'deterioration', 'emergency' and 'department' in all fields; with any of '"track and trigger"', '"early warning"', '', 'OR', 'EWS', 'OR', 'MEWS', 'OR' or 'POTTS' in all fields	0
		All of 'patient', 'deterioration', 'emergency' and 'department' in all fields	1
		All of 'early', 'warning', 'emergency' and 'department' in all fields	0
		All of 'early' and 'warning' in all fields	37. After sifting = 4
		All of 'track', 'and' and 'trigger' in all fields	2. After sifting = 1
Proquest Dissertations and Theses UK & Ireland	www.library.nuigalway.ie	ab(emergency OR trauma OR casualty OR triage) AND (care OR ward* OR department* OR unit* OR room* OR health* OR hospital OR service) AND (warning OR trigger OR EWS OR MEWS OR POTTS) AND (deteriorate* OR worse* OR adverse OR weaken* OR weaker OR "acute illness" OR Monitor* OR escalate* OR reassess*) AND adult*	24 after sifting = 0
Proquest Dissertations and Theses A & I	www.library.nuigalway.ie	ab(warning OR trigger OR EWS OR MEWS OR POTTS) AND ab(deteriorate* OR worse* OR adverse OR weaken* OR weaker OR "acute illness" OR Monitor* OR escalate* OR reassess*) AND adult*	23. After sifting = 0
		ab(emergency OR trauma OR casualty OR triage) AND ab(care OR ward* OR department* OR unit* OR room* OR health* OR hospital OR service) AND ab(warning OR trigger OR EWS OR MEWS OR POTTS) AND (deteriorate* OR worse* OR adverse OR weaken* OR weaker OR "acute illness" OR Monitor* OR escalate* OR reassess*) AND adult*	88. After sifting = 0

Clinical Trials Registries			
Date of Search: 12-13/03/2016			
Website/ Database	**URL**	**Search Terms**	**Hits**
CENTRAL	www.cochranelibrary.com	See Cochrane Library tab	n/a
Prospero	www.crd.york.ac.uk/PROSPERO	Track and trigger	3. After sifting = 0
		Early warning	12. After sifting = 2
		Patient deterioration emergency department	0
		Patient deterioration emergency	0
		Patient deterioration	5. After sifting = 0
		EWS	3. 1 relevant but already picked up in early warning search. 0
		MEWS	6. 1 relevant but already picked up in early warning search. 0
		POTTS	4. After sifting 0
ClinicalTrials.gov	https://clinicaltrials.gov/ct2/search/advanced	"patient deterioration" AND (emergency OR trauma OR casualty OR triage) AND (warning OR trigger OR escalation OR EWS OR MEWS OR POTTS) Filters used: Adult (18–65) & Senior (66+)	0
		patient deterioration AND (emergency OR trauma OR casualty OR triage) AND (warning OR trigger OR escalation OR EWS OR MEWS OR POTTS) Filters used: Adult (18–65) & Senior (66+)	9. After sifting = 1
		(emergency department OR trauma OR casualty OR triage) AND (early warning OR track and trigger OR escalation OR EWS OR MEWS OR POTTS) Filters used: Adult (18–65) & Senior (66+)	107. After sifting = 2

World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP)	http://apps.who.int/trialsearch/AdvSearch.aspx	Advanced search strategy. Only options to limit to 1) title 2) condition or 3) intervention. Tried using following in intervention field: early warning or track and trigger or ews or mews or potts.	1138 results. Skim through shows low specificity therefore strategy abandoned.
		Advanced search option not useful. Basic search : (emergency OR trauma OR casualty OR triage) AND (care OR ward* OR department* OR unit* OR room* OR health* OR hospital OR service) AND (warning OR trigger OR escalate* OR EWS OR MEWS OR POTTS) AND (deteriorate* OR worse* OR adverse OR weaken* OR weaker OR "acute illness") (Monitor* OR escalate* OR reassess*) AND (adult*)	Site unable to handle this strategy and kept crashing
		Basic option used again: patient deterioration AND emergency department OR Triage AND early warning OR track and trigger OR EWS OR MEWS OR POTTS	29. After sifting 9 were useful but overlap with Clinicaltrials.gov results. 6 kept

Appendix 2: Checklists for rapid reviews (Abrami et al 2010)[31]

Key aspects	Considerations/recommendations	Section where it has been addressed
Research question	The research question should be clearly stated and any decision to limit the treatment conditions and outcomes measures should be explained.	2
Inclusion criteria	An explanation of the inclusion and exclusion criteria should be provided, noting whether the review will be limited by time, source, location, context or methodological quality.	3.1
Search strategies	The electronic and other search strategies should be explicitly stated so that it is clear how the review is comprehensive and free of bias.	3.2
Inter-rater agreement	The review should describe the steps taken to ensure inter-rater agreement during the phases of study identification, calculation of effects and coding of study features (if applicable).	3.3-3.5 (independent screening and quality assessment)
Effect extraction	The method used to aggregate results should be explained and defended.	3.5
Study features	The review should describe whether or how variability among studies was explored, and if not, why not.	4
Analysis	Measures of central tendency and variability should be reported. The heterogeneity of effects should be noted.	N/A
Interpretation and implications	The review should contain a clear set of conclusions and implications.	5
Cautions and limitations	The conclusion should outline in the ways in which the brief review differs from a comprehensive review, especially the limitations of brief review methodology and the risk associated with the truthfulness of the findings.	5
Other	Sources of evidence (i.e. publication information) should be available but not necessarily included in the report.	References

Appendix 3: Risk of bias and methodological quality assessment

Descriptive studies: Extent of use

Assessed using the Quality Assessment Tool adapted from NIH (2014)[23]

Note: the studies shown here relate to table 4 of the evidence tables.

Criteria	Australian Commission (2011)[47]	Considine et al (2012)[34]	Correia et al (2014)[35]	Coughlan et al (2015)[33] Conference abstract	Griffiths et al (2012)[36]	Wilson et al (2013)[37]
1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes	No	Yes	Yes
3. Was the participation rate at least 50%?	Yes	No (Total sample represented 9.8% of ED EWS activations)	Yes	NR	Yes (57% response rate)	Yes
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study pre-specified and applied uniformly to all participants?	Yes	Yes	Yes	NR	NR	No (Restricted to times when a member of research team was on duty)
5. Was a sample size justified?	No	No	No	NR	No	No
6. Were the outcome measures clearly defined, valid, reliable, and implemented consistently across all study participants?	NR	Yes	Yes	NR	Yes	Yes
Quality Rating (Good, Fair, or Poor)	**Fair**	**Fair**	**Fair**	**Poor**	**Fair**	**Fair**

*CD, cannot determine; NA, not applicable; NR, not reported

Descriptive studies: Compliance

Assessed using the Quality Assessment Tool for Before-After (Pre-Post) Studies with No Control Group (NIH 2014)[23]

Note: the studies shown here relate to table 5 of the evidence tables.

Criteria	Austen et al (2012)[40]	Christensen et al (2011)[38]	Hudson et al (2015)[41]	Johnson et al (2014)[39]
1. Was the study question or objective clearly stated?	Yes	Yes	Yes	Yes
2. Were eligibility/selection criteria for the study population pre-specified and clearly described?	Yes	Yes	No	Yes
3. Were the participants in the study representative of those who would be eligible for the test/service/intervention in the general or clinical population of interest?	Yes	Yes	Yes	Yes
4. Were all eligible participants that met the pre-specified entry criteria enrolled?	Yes	Yes	CD	Yes
5. Was the sample size sufficiently large to provide confidence in the findings?	CD	Yes	CD	Yes
6. Was the test/service/intervention clearly described and delivered consistently across the study population?	Yes	Yes	Yes (post)	N/A
7. Were the outcome measures pre-specified, clearly defined, valid, reliable, and assessed consistently across all study participants?	Yes	Yes	No (not pre-specified)	Yes
8. Were the people assessing the outcomes blinded to the participants' exposures/ interventions?	N/A	N/A	NR	N/A
9. Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?	Yes	Yes	Yes	Yes
10. Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes?	N/A	N/A	Yes	N/A
11. Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention (i.e., did they use an interrupted time-series design)?	N/A	N/A	N/A	N/A
12. If the intervention was conducted at a group level (e.g., a whole hospital, a community, etc.) did the statistical analysis take into account the use of individual-level data to determine effects at the group level?	N/A	N/A	N/A	N/A
Quality Rating (Good, Fair, or Poor)	**Good**	**Good**	**Fair**	**Good**

*CD, cannot determine; NA, not applicable; NR, not reported

Guidelines:

Assessed using the AGREE II quality assessment tool[24]

Note: the studies shown here relate to table 6 of the evidence tables.

AGREE II TOOL (2 reviewers)			
	NCG No. 1 NEWS[7]	**NICE CG50[9]**	**RoCP NEWS[8]**
DOMAIN 1: Scope and purpose (3 items)	97.2%	91.7%	91.7%
DOMAIN 2: Stakeholder involvement (3 items)	91.7%	91.7%	91.7%
DOMAIN 3: Rigour of development (8 items)	93.8%	92.7%	69.8%
DOMAIN 4: Clarity of presentation (3 items)	91.7%	100.0%	100.0%
DOMAIN 5: Applicability (4 items)	91.7%	89.6%	64.6%
DOMAIN 6: Editorial independence (2 items)	95.8%	87.5%	62.5%
OVERALL GUIDELINE ASSESSMENT (1 items)	91.7%	91.7%	66.7%
I would recommend this guideline for use.	Yes	Yes	Yes with modification

Effectiveness studies:

Assessed using the EPOC Risk of Bias Tool[27]

Note: the studies shown here relate to table 7 of the evidence tables.

EPOC criteria	Shuk-Ngor et al (2015)[42] Two group non-randomised comparison (MEWS group versus usual observation group)
Was the allocation sequence adequately generated?	High risk of bias
Was the allocation adequately concealed?	High risk of bias
Were baseline outcome measurements similar?	Unclear risk of bias
Were baseline characteristics similar?	Low risk of bias
Were incomplete outcome data adequately addressed?	Low risk of bias
Was knowledge of the allocated interventions adequately prevented during the study?	Unclear risk of bias
Was the study adequately protected against contamination?	Unclear risk of bias
Was the study free from selective outcome reporting?	Low risk of bias
Was the study free from other risks of bias?	Unclear risk of bias
Overall Risk of Bias Judgement	**High risk of bias**

Effectiveness studies:

Assessed using the GRADE assessment of quality of evidence[26]

Note: the studies shown here relate to table 7 of the evidence tables.

Study: Shuk-Ngor et al (2015)[42]														
Outcome	**No. of studies**	**Design**	**ROB**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**	**Large magnitude effect**	**Dose-response gradient**	**Effect of plausible residual confounding**	**MEWS**	**Usual observation**	**Relative risk**	**Quality**
Change in management (no of activations)	1	Non-RCT	Serious ROB[a]	No serious inconsistency	No serious indirectness	Serious imprecision[b]	No serious publication bias	No large effect	N/A	N/A	1/10 of 269	1/20 of 275	2.0 (95% CI 1.1; 3.8)	Very low
Adverse events	1	Non-RCT	Serious ROB[a]	No serious inconsistency	No serious indirectness	Serious imprecision[b]	No serious publication bias	No large effect	N/A	N/A	1/269	1/275	1.02 (95% CI 0.06; 16.3)	Very low

[a] Downgraded one level because the ROB for this study was rated as high.
[b] Downgraded one level because of small number of events in the study.

Development and validation studies – Scoping Review:

Assessed using the Quality Assessment Tool adapted from NIH (2014)[23]

Note: the studies shown here relate to table 9 of the evidence tables.

Criteria (Quality Assessment Tool adapted from NIH (2014))	Challen et al (2011)[46] A Scoping review
1. Was the research question or objective in this paper clearly stated?	Yes
2. Was the study population clearly specified and defined?	N/A
3. Was the participation rate at least 50%?	N/A
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study pre-specified and applied uniformly to all participants?	Yes
5. Was a sample size justified?	N/A
6. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes
Quality Rating (Good, Fair, or Poor)	**Good**

Development and validation studies:

Assessed using the Quality Assessment Tool (adapted from Kansagara *et al* 2011)

Note: the studies shown here relate to tables 10, 11 and 12 of the evidence tables and are presented alphabetically by author.

Study (year)	Adequate description of population[a]	Non-biased selectionb	Adequate prognostic factor measurementc	Adequate outcome measurementd	Method of validatione	Overall risk of bias
Albright et al (2014)53 (Development & Validation)	Yes	No (Only those who had blood cultures or influenza swab were included)	Yes (SOS (for high risk of outcome) ≥ 6)	Yes	Yes (New instrument against other ED system; AUROC and SOS ≥ 6 versus < 6)	Low risk of bias
Alam et al (2015)76 (Validation)	Yes	No (Recruitment between 12-8pm only)	Yes (NEWS 0-4, 5-6, ≥7)	Yes	Unclear (AUROC reported for only 2 of 4 outcomes)	Unclear risk of bias
Armagan et al (2008)58 (Validation)	Yes	Unclear (No clear statement)	Yes (mEWS > 4)	Yes	Unclear (Multi-variate regression only)	Unclear risk of bias
Bulut et al (2014)59 (Validation)	Yes	Unclear (No clear statement)	Yes (MEWS ≥ 5 REMS > 13)	Yes	Yes (Against other system; AUROC)	Low risk of bias
Burch et al (2008)45 (Validation)	Yes	No (Every 6th day only and 790 patients included = 70.2% of potential study cohort)	Unclear (Indicates MEWS ≥ 5 but no clear statement of cut-off for high risk of outcome)	Yes	Unclear (Univariate regression only)	High risk of bias
Cattermole et al (2009)61 (Development & Validation)	Yes	Yes	Unclear (No clear statement of cut-off scores for systems)	Yes	Yes Against other systems; multivariate regression and AUROC	Low risk of bias

Study (year)	Adequate description of population[a]	Non-biased selectionb	Adequate prognostic factor measurementc	Adequate outcome measurementd	Method of validatione	Overall risk of bias
Cattermole et al (2014)62 (Development & Validation)	Yes	No (Week days only)	Unclear (No clear statement of cut-off scores for systems)	Yes	Yes (Against other systems; multivariate regression and AUROC)	Unclear risk of bias
Christensen et al (2011)69 (Development & Validation)	Yes	Yes	Yes BEWS ≥ 5	Yes	Yes (BEWS ≥ 5 versus BEWS < 5; sensitivity, specificity)	Low risk of bias
Cıldır et al (2010)60 (Validation)	Yes	Unclear (No clear statement)	Yes (Sepsis defined MEWS > 4 CCI > 4)	Yes	Yes (Against other systems; survivor versus non-survivor and AUROC)	Low risk of bias
Considine et al (2015)78 (Validation)	Yes	Yes (stratified random sample)	Yes (Single parameter system)	Yes	No (Only Mann-Whitney U and Kruskal Wallis tests)	Low risk of bias
Corfield et al (2014)81 (Validation)	Yes	Yes	Yes (N/A for high risk prediction cut-off scores as development study)	Yes	Yes (Regression; survivor versus non-survivor and AUROC)	Low risk of bias
Dundar et al (2015)79 (Validation)	Yes	Yes	Yes (optimal cut-off determined by Youden's index)	Yes	Yes (Two systems compared and AUROC)	Low risk of bias

Study (year)	Adequate description of population[a]	Non-biased selectionb	Adequate prognostic factor measurementc	Adequate outcome measurementd	Method of validatione	Overall risk of bias
Eick et al (2015)70 (Validation)	Yes	Yes	Unclear Cut-offs for DC and MEWS (high-risk) not clearly provided	Yes	Yes (Against other systems; multivariate regression; bootstrapping and AUROC)	Low risk of bias
Geier et al (2013)71 (Development & Validation)	Yes	Yes	Yes (ESI ≤ 2; MEWS ≥ 5 MEDS ≥ 8; CCI ≥ 2)	Yes	Yes (Against other systems and AUROC)	Low risk of bias
Graham et al (2007)63 (Conference abstract only) (Validation)	Yes	Yes	Unclear (Insufficient detail)	Yes	Unclear (Mentions AUROC but insufficient detail to fully assess)	Unclear risk of bias
Gu et al (2015)32 (Only abstract in English) (Validation)	Unclear Partial details given in results but no clear statement on selection criteria	Unclear No clear statement	Yes (MEWS ≥ 5)	Yes	Yes (Multivariate regression and MEWS positive (≥ 5) versus MEWS negative (0-4))	Unclear risk of bias
Heitz et al (2010)54 (Validation)	Yes	Yes	Yes (N/A for high risk prediction cut-off scores as development study)	Yes	Yes (Multivariate regression; AUROC and MEWS Max cut-of scores (≥ 1 through to ≥ 9))	Low risk of bias

Study (year)	Adequate description of population[a]	Non-biased selectionb	Adequate prognostic factor measurementc	Adequate outcome measurementd	Method of validatione	Overall risk of bias
Ho et al (2013)64 (Validation)	Yes	No (No clear statement and 8am to 6pm recruitment only)	Yes (MEWS ≥ 4)	Yes	Yes (Regression; AUROC and MEWS < 4 versus MEWS ≥ 4)	Low risk of bias
Hock Ong et al (2012)65 (Validation)	Yes	No (No clear statement and 'office hours' recruitment only)	Unclear (Cut-offs for DC and MEWS (high-risk) not clearly provided)	Yes	Yes (MEWS versus ML system and AUROC)	Unclear risk of bias
Howell et al (2007)55 (Validation)	Yes	Yes	Unclear (Refers to other publication for calculations)	Yes	Yes (Against other systems; regression and AUROC)	Low risk of bias
Jo et al (2013)73 (Validation)	Yes	Yes	Unclear (Cut-off for VIEWS-L not clearly provided)	Yes	Yes (Against other system; AUROC)	Low risk of bias
Jo et al (2016)72 (Validation)	Yes	Unclear	Yes (Optimal cut-off determined by Youden Index)	Yes	Yes (Against other systems; AUROC)	Low risk of bias
Jones et al (2005)56 (Validation)	Yes	Unclear (No clear statement)	Yes (Predicted mortality > 50%)	Yes	Yes (Against other systems; AUROC)	Low risk of bias
Junhasavasdikul et al (2013)74 (Validation)	Yes	Unclear (States all but no clear statement if consecutive enrolment occurred)	Unclear (Vital sign cut-off values for high-risk not provided)	Yes	Yes (Multivariate regression and R2)	Unclear risk of bias

Study (year)	Adequate description of population[a]	Non-biased selectionb	Adequate prognostic factor measurementc	Adequate outcome measurementd	Method of validatione	Overall risk of bias
Keep et al (2016)50 (Validation)	Yes	Yes	Unclear (Cut-off values for high-risk not provided)	Yes	Yes (NEWS cut-off ≥ 1 through to ≥ 11; AUROC)	Low risk of bias
Liu et al (2014)66 (Development & Validation)	Yes	Unclear (No clear statement)	Unclear (No detail on scores)	Yes	Yes (Against other systems and AUROC)	Unclear risk of bias
Naidoo et al (2014)44 (Validation)	No (Selection criteria unclear)	No (Every 5th record)	Yes (TEWS ≥ 7)	Yes	Unclear (TEWS < 7 versus TEWS ≥ 7 only)	High risk of bias
Nguyen et al (2012)57 (Validation)	Yes	Unclear (Only patients enrolled in the registry for whom all 3 physiologic scores were available)	Yes	No (Reports mortality but it is not pre-specified or defined)	Yes (Against other systems; AUROC)	Unclear risk of bias
Olsson & Lind (2003)67 (Development & Validation)	Yes	Yes	Unclear (No clear statement on cut-off scores for high risk)	Yes	Yes (Split sample technique; multivariate regression; AUROC)	Low risk of bias
Olsson et al (2004)68 (Development & Validation)	Yes	Yes	Unclear (No clear statement on cut-off scores for high risk)	Yes	Yes (Split sample technique; multivariate regression and AUROC)	Low risk of bias

Study (year)	Adequate description of population[a]	Non-biased selectionb	Adequate prognostic factor measurementc	Adequate outcome measurementd	Method of validatione	Overall risk of bias
Subbe et al (2006)51 (Validation)	Yes	Yes	Yes (Cut-off scores for risk prediction provided for each system)	Unclear (ICU admissions reported but outcome not pre-specified or defined)	Yes (Against other systems (correlations))	Low risk of bias
Vorwerk et al (2009)52 (Validation)	Yes	Yes	Yes (Cut-off scores for risk prediction provided for each system)	Yes	Yes (By outcome groups and AUROC)	Low risk of bias
Wang et al (2016)75 (Validation)	Yes	Unclear (Only 99 of 234 initially eligible had a peri-arrest MEWS)	Unclear (No clear statement on cut-off scores for high risk)	Yes	Unclear (Multivariate regression only)	Unclear risk of bias
Williams et al (2016)77 (Validation)	Yes	Yes	Yes (Table 1; score > 15 = 50% predicted mortality)	Yes	Yes (Against other systems and AUROC)	Low risk of bias
Wilson et al (2016)43 (Validation)	Unclear (Selection criteria unclear)	No (Recruitment restricted to times when research team available)	Unclear (Cut-off values not provided)	Yes	Unclear	High risk of bias

[a] Study describes inclusion criteria for selecting patients, and for enrolled patients describes duration and severity of symptoms, demographics (at least age), and setting (primary care vs. occupational vs. other).

[b] Study either reports enrolling (or attempting to enrol) a consecutive series of patients meeting inclusion criteria, or a random sample.

[c] Study describes reproducible and appropriate methods for measuring prognostic factors

[d] Study describes reproducible and appropriate methods to define and identify outcome

[e] Method of validation is clear and appears to be appropriate

An Roinn Sláinte
Department of Health

Department of Health, Hawkins House, Hawkins Street, Dublin 2, D02 VW90, Ireland
Tel: +353 1 6354000 • Fax: +353 1 6354001 • www.health.gov.ie